Essays on Credit Supply and Demand in the Housing Boom and Bust

by

Daniel García

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
June, 2018

Abstract

The housing boom of the mid 2000s proved to be unsustainable, and led to the Great Recession of 2007-2009 – the worst economic collapse in the United States since the Great Depression of the 1930s. This dissertation empirically assesses some of the main explanations for the severity of the housing boom and bust.

In the first chapter, "Employment in the Great Recession: How Important were Household Credit Supply Shocks?," I pool data from all large multimarket lenders in the U.S. to estimate how many of the seven million jobs lost in the Great Recession can be explained by reductions in the supply of mortgage credit. I construct a mortgage credit supply instrument at the county level, as the weighted average (by prerecession mortgage market shares) of liquidity-driven lender shocks during the recession. I find that the reduction in mortgage supply could explain about 15 percent of the employment decline. The job losses are concentrated in construction and finance.

In the second chapter, "Property Investors and the Housing Boom and Bust," I use new cross-sectional housing data to argue that property investment – existing homeowners acquiring additional properties – was a central driver of boom-bust dynamics over the recent housing cycle. Measuring investor activity at the county level as the fraction of mortgage originations for non-owner-occupied housing, I find that 'investor' counties with high amenity values (warm winters, waterfronts) had high investor activity both before the 2000s and in the peak boom years. In counties with high investor shares in 1998-2000, home prices and employment grew faster in 2003-2006 than elsewhere, and crashed harder in 2007-2010. My estimate is that investor activity could explain

30 percent of the total variation in construction and financial employment over 2003-2010.

High savings in emerging economies may have helped fuel the U.S. housing boom. To shed light on why rapidly urbanizing high growth countries such as China are capital exporters, the third chapter "Can Risky Rural-Urban Migration Help Explain the Flow of Capital from Developing to Advanced Economies?" models the saving motives of residents in a country undergoing rapid urbanization characterized by circular migration, a feature of many developing nations. Migrants move back and forth between rural-urban areas as determined by transition probabilities calibrated to match migration flows and the pace of urbanization in China. Workers accumulate savings in the high-productivity urban sector because of the risk of returning to the low-wage rural sector. As the urban population increases, so do aggregate savings and income.

Advisors: Christopher D. Carroll, Jon Faust, Jonathan Wright

Acknowledgement

This dissertation would not have been possible without the help of many advisors, colleagues, friends, and family.

I am deeply grateful to my advisors Chris Carroll, Jon Faust, and Jonathan Wright for their guidance, support, and encouragement. Writing papers requires not just academic rigor, but also vision, and I am lucky to have had such a strong team of advisors to learn from. I am also grateful to Olivier Jeanne, Laurence Ball, Robert Barbera, Greg Duffee, and Anton Korinek for their mentorship. It takes a village to raise an economist. Thanks as well to Alina Arefeva, Lingxin Hao, and Benjamin Hobbs for agreeing to be part of the dissertation committee.

Work on the first and second chapters of the dissertation was funded by fellowships at the Federal Reserve Board, Bank of Chicago, and the AEA's CSWEP-CSMGEP. I am very grateful for this support. Spending time at the Fed was easily one of the highlights of my Ph.D. Special thanks are due to Raven Molloy, Steven Laufer, Neil Bhutta, Paul Lengermann, Laura Feiveson, Andrew Paciorek, Aditya Aladangady, John Driscoll, Alejandro Justiniano, Egon Zakrajsek, and John Rogers.

I would also like to thank the staff and students at Hopkins for their support and friendship, especially Ryan Bush, Jiaxiong Yao, Sohini Mahapatra, Burcin Kisacikoglu, Alanna Borklund, Victor Ronda, Emmanuel García, and Chang Ma. Thanks, too, to my friends (shoutout to Grinnell) – always an invaluable source of support, fun, and help keeping it real.

Last, but not least, I could not have done this without my family – my brothers and their families, and my parents in Honduras.

Contents

List of Tables

List of Figures

Chapter 1

Employment in the Great Recession: How Important were Household Credit Supply Shocks?

1.1 Introduction

Employment fell by over 7 million in the Great Recession. Possible explanations include declines in credit supply (Eggertsson and Krugman 2012; Guerrieri and Lorenzoni 2011), household net worth (Mian and Sufi 2014; Giroud and Mueller 2015), and increases in uncertainty (Baker, Bloom, and Davis 2015; Bloom 2014). The goal of this paper is to isolate and empirically assess the credit supply hypothesis: to what extent did reductions in credit supply play a causal and independent role in explaining the job losses that occurred in the period 2007-2010? To do so, I measure plausibly exogenous variation in credit supply (specifically for mortgages) at the county level, based on the interaction of prerecession county-lender market shares, which measure the importance of each lender to each locality immediately prior to the recession, and heterogeneous aggregate lender shocks during the recession. The county level estimates show that the reduction in mortgage supply negatively affected the health of residential markets, leading to declines in home buying, home prices, and employment in the construction sector; in other industries less directly linked to real estate, the job losses were much more muted and close to zero. A partial equilibrium aggregation exercise suggests that the decline in mortgage supply could explain close to 15 percent of the total job losses during the recession, or about 1 million of the total jobs lost.

The starting point of this paper is the observation that there is a strong OLS association between declines in local employment and mortgage credit issuance during the recession. This suggests that reductions in mortgage supply could have played an important role in driving the job losses. On the other hand, the OLS association could be entirely driven by reverse causality – declines in local employment and economic activity could have led to the decline in mortgage issuance. To isolate the effects of reductions in mortgage supply on economic activity, I construct a mortgage credit supply instrument at the county level.

The identification strategy exploits the well-known fact in the mortgage literature (discussed and further documented in the paper) that credit relationships in the mortgage market – as in the corporate market – are persistent and not easily substitutable.[1] Therefore, an exogenous lender shock to a locality is not

[1]Market shares at the county-lender level are highly persistent year on year; for example, 2005 county-lender shares explain 2007 shares almost 1-for-1. In the recession, there were few cases of lender entry into new localities: of 2008-2010 county-lender pairs, less than 8 percent were new. Even in 'normal' times there is limited shopping, with Alexandrov and Koulayev (2017), Woodward and Hall (2012) and Lacko and Pappalardo (2010) finding that borrowers shop too little despite significant price dispersion.

immediately offset by new lender entry.

The instrument measures the average supply response of a county's traditional lenders during the recession for 'nonlocal reasons' – reasons unrelated to the condition of local economies. The instrument is based on two sources of variation: (i) the heterogenous aggregate supply response of lenders during the recession, and (ii) variation in the reliance of localities to different lenders prior to the recession (measured with 2005-2007 market shares). To measure (i) aggregate differences in lender supply, I estimate lender fixed effects explaining variation in credit changes at the county-lender level during the recession, while holding constant local economic conditions (via county fixed effects). The lender fixed effects estimates are highly robust to alternative specifications, such as controlling for census tract fixed effects or loan characteristics varying at the county-lender level.[2] County-lender market shares (ii) come straight from the main data source, the Home Mortgage Disclosure Act. The county level credit supply instrument is the weighted average (by 2005-2007 market shares) of the lender fixed effects.

This paper is the first to construct a Bartik-style instrument based on the interaction of heterogeneous aggregate lender shocks and local market shares in the mortgage market during the recession. Working in parallel, Mondragon (2018) also studies the employment effects of household credit shocks during the recession, though his main instrument is based on county exposure to a single troubled lender during the recession (discussed shortly). My approach

[2]Specifically, I regress credit changes at the county-lender level over 2007-2010 on county fixed effects and lender fixed effects in the baseline specification. The lender fixed effects are highly correlated (close to 1) when controlling instead for census tract fixed effects, using only high-income or low-income loans, and controlling for variation in precrisis county-lender loan characteristics.

instead pools data from essentially all large multimarket lenders in the U.S., and follows in the tradition of recent related work studying the employment effects of reductions in *corporate* credit supply via Bartik-style instruments, such as Chodorow-Reich (2014), Greenstone, Mas, and Nguyen (2015), and Amiti and Weinstein (2018). The main contribution of this paper is the focus on mortgage supply during the crisis, which may be particularly important, since mortgages are the largest category of private credit, and funding markets for mortgages were severely disrupted during the crisis. Notably, the private secondary market for mortgages fully collapsed at the onset of the crisis and remained inactive throughout. In line with previous research documenting that low liquidity contributed to lower credit issuance during the crisis (Ivashina and Scharfstein 2010; Cornett et al. 2011; Irani and Meisenzahl 2015), I document that mortgage lenders were more likely to cut supply if they relied on funding sources that proved fragile in the crisis. In particular, the reliance of banks on wholesale debt, loan sales in the secondary market, and especially loan sales to private buyers, explains 72 percent of the variation in lender supply during the recession (the lender fixed effects).

The 2SLS results controlling for a detailed set of county observables and region fixed effects are as follows. Declines in mortgage supply negatively affected the health of residential markets. For example, a supply-driven plausibly exogenous 10 percent decline in local mortgage issuance led to a 10 percent decline in new residential permits and a 5 percent decline in home prices. Areas with larger declines in mortgage supply also experienced higher default and foreclosure rates. The next question is whether the negative shock to real estate spilled over into local labor markets, both in directly related industries such as construction and in other industries.

The employment effects are largely direct and concentrated in construction and finance, a category of employment where about a third of workers are real estate intermediaries.[3] The main mechanism is that declines in mortgage supply reduce demand for housing, which contributes to job losses in industries reliant on housing demand. As evidence, I find that, for a given decline in mortgage credit, job losses in construction are larger in counties where housing supply is more elastic – areas where construction is more responsive to changes in housing demand. The estimated effects on other categories of employment – total private employment excluding construction and finance, and nontradable employment – are close to zero and not significant. The 2SLS estimates on these broad employment categories contrast with their OLS counterparts, which are 2-3 times larger and highly significant, suggesting the OLS estimates are biased upward due to reverse causality.

Overall, a 10 percent exogenous decline in local mortgage credit leads to a 1 percent decline in total private employment. To gain a sense of the aggregate implications of the county-level estimates, I perform a partial equilibrium aggregation exercise that exploits the in-sample distribution of the credit supply instrument, similar to the approaches in Chodorow-Reich (2014) and Mian and Sufi (2014). The exercise suggests reductions in supply could explain close to 15 percent of the total jobs lost, or about 1.1 million jobs. The bottom line is that the reduction in mortgage supply likely aggravated the fall in employment to a meaningful, but moderate, extent.

[3]Construction and finance accounted for close to 35 percent of the job losses in the recession. Typically, their share in total employment is between 10-15 percent.

The main concern regarding instrument validity is that lender location might be correlated with unobserved local characteristics associated with job losses during the recession. Identification requires that below-average suppliers were not systematically sorted into localities experiencing below-average (or above-average) employment shocks. To the extent that counties and lenders are matched along observables, controlling for those characteristics isolates the remaining 'as good as random' variation in lender location. To that end, I control for a highly detailed set of local characteristics that explain about 60 percent of the variation in mortgage credit issuance across localities during the recession, including the share of subprime borrowers, the run-up in home prices during the boom, and various other demographic, housing, and industry characteristics. It is not possible to control for everything that may be relevant, however, and so I also rule out specific hypotheses about non-random lender location. For example, risky lenders may have moved to risky localities during the boom years. However, measuring the exposure of lenders to localities using 2000-2002 (instead of 2005-2007) market shares yields very similar results – the first stage is weaker due to the loss in precision, but 2SLS point estimates are not statistically different. I also show that results are very similar when using region, division, or state fixed effects – this rules out hypotheses such as the possibility that weak suppliers in the recession were more heavily concentrated in the Sand States.[4]

This paper is part of the literature exploring the extent to which credit shocks explain the fall in employment in the Great Recession. Most empirical

[4]The results are robust to a number of checks including 'placebo' tests; controlling for the decline in small business lending in the recession; and the inclusion of large failed lenders (e.g., IndyMac) in the analysis.

work focuses on the employment effects of corporate credit shocks.[5] Chodorow-Reich (2014) estimates that credit shocks explain between one-third and one-half of job losses at firms borrowing in the syndicated loan market. Greenstone, Mas, and Nguyen (2015) find that credit shocks to small business loans help explain declines in borrowing, but produce only small employment effects. The credit supply variation I measure is specific to mortgages because county-lender market shares for mortgages and small business loans are largely uncorrelated – in other words, the mortgage lenders to one locality are often not the same as the small business loan lenders. The results in the paper are robust to controlling for declines in small business loans.

The most closely related paper is Mondragon (2018), whose credit supply instrument is exposure to Wachovia Bank, a troubled lender in the recession. He estimates that a 10 percent decline in instrumented mortgage credit leads to a 3 percent decline in employment, an elasticity two times as large as the OLS counterpart, and three times as large as my own estimate. One concern is potential 'bad' bank in a 'bad' region matching – Wachovia had a larger presence in states in the South Atlantic (e.g., FL, SC, NC) where job losses were among the worst in the country. His estimates would have an upward bias if employment shocks and Wachovia location have correlated spatial fixed effects. In fact, when using division or state fixed effects, the Wachovia instrument significantly weakens. In contrast, my paper pools information from all large lenders located across the U.S. and employs a richer set of county controls, and so is more robust to potential concerns about non-random county-lender

[5] A related empirical literature studies the international transmission of the financial crisis through the banking sector (Cetorelli and Goldberg (2011); Haas and Lelyveld (2014); Schnabl (2012)).

matching. For example, the 2SLS point estimates in this paper are essentially the same (not statistically different) when using region, division, or state fixed effects.

The paper is also related to the work of Mian and Sufi (2014), Mian, Rao, and Sufi (2013), Kaplan, Mitman, and Violante (2017), and others, on the household net worth channel, which assesses the hypothesis that declines in household net worth led to declines in aggregate demand and employment. The credit supply and household net worth channels are related – for example, the bursting of the housing boom helped precipitate the financial crisis. However, the run-up in house prices does not explain all of the ensuing economic decline. To isolate the credit supply channel, this paper asks: holding house prices constant, what were the employment effects of reductions in mortgage supply during the recession? I therefore condition on house price changes during the boom years, as well as various prerecession characteristics of localities associated with the housing boom and bust.

More broadly, this paper is part of the literature studying the effects of changes in mortgage supply on housing and labor markets. Most empirical work has focused on the former. Related work includes Favara and Imbs (2015); Mian and Sufi (2011); Adelino, Schoar, and Severino (2012); Berrospide, Black, and Keeton (2016); Glancy (2015); Anenberg et al. (2016); Vojtech, Kay, and Driscoll (2016); Gropp, Krainer, and Laderman (2014); Gete and Reher (2016); Haltenhof, Jung Lee, and Stebunovs (2014); Chen, Hanson, and Stein (2017). Only a handful of papers focus on the employment effects of reductions in mortgage supply. In the boom years, DiMaggio and Kermani (2014) use a federal preemption of national banks from local anti-predatory lending laws in 2004 to estimate the elasticity of nontradable employment with respect to mortgage supply. In the bust, Passmore and Sherlund (2016) find that counties more reliant on GSEs for mortgage credit experienced healthier labor markets in the Great Recession. I contribute to this literature by highlighting the heterogeneous industry effects of mortgage supply shocks on construction and financial employment.

1.2 Data Sources

I assemble a detailed county-level dataset including home prices, home sales, employment, mortgage credit, credit scores, demographics, borrower characteristics, industry composition, and various other local characteristics. The main source for mortgage data is the Home Mortgage Disclosure Act (HMDA). Mortgage lenders with offices in metropolitan areas are required to publicly disclose detailed information each year, including the dollar amount and number of mortgages issued, as well as the location (census tract, county) of the property securing the loan. Throughout the mid to late 2000s, HMDA covered over 90% of residential mortgage lending by dollar amount (Dell′Ariccia, Igan, and Laeven 2012). I use mortgages for home purchase and improvement as the main measure (loan purpose 1 and 2 in HMDA). Figure 1.1 plots aggregate trends in mortgage originations, total private employment, and the S&P Case-Schiller U.S. National Home Price Index, with the series indexed to their 2006 value.

Data on delinquency rates, foreclosure rates, home sales, and home prices are obtained from CoreLogic. Data on building permits comes from the Census. For employment, I rely on two sources, both of which are establishment-based and provide nearly full coverage of private employment: the Quarterly Census of Employment and Wages (QCEW), and the County Business Patterns (CBP). I use the CBP to measure tradable and nontradable employment using the definitions in Mian and Sufi (2014), and the QCEW for the other employment data.

Table 1.1 shows summary statistics for over 1,000 of the largest counties in the U.S. Each of these localities had over 15,000 households in the 2000 Decennial Census and account for almost 85% of aggregate employment. Table 1.3 provides definitions and sources for the data used throughout the paper. While mortgage credit declined over 2007-2010 in virtually all counties, there is significant cross-sectional variation in the decline, with credit falling by more than 51% in ten percent of the counties in the sample and falling by less than 21% in the top decile. Figure 1.2 shows there is a strong positive OLS association between declines in mortgage credit issuance and declines in both home prices and employment. This suggests that declines in mortgage issuance could have driven employment losses. On the other hand, the relationship might be entirely explained by reverse causality – declines in local economic activity could have driven the decline in employment and credit issuance.

I obtain lender-level data from HMDA, which provides loans by lender subsidiaries (respondents) and locality. I match subsidiaries belonging to the same parent company using the crosswalk maintained by Robert Avery, and aggregate to the level of the parent company (bank holding company, for banking institutions).[6] To calculate changes in lending at the lender level without including changes due to acquisitions, I use the standard approach (Bernanke and

[6] Available upon request at Robert.Avery@fhfa.gov

Lown 1991; Greenstone, Mas, and Nguyen 2015) of treating the acquired and acquiring institutions as part of the same entity throughout the sample period, which in this paper is over 2000-2010. I also conservatively drop failed institutions for most of the paper, because the extent to which their credit decline was supply- or demand-driven cannot be credibly estimated. Dropping these institutions is a conservative choice: it reduces the potential for biased estimates at the expense of statistical power. I show, however, that including the failed lenders, by assuming all of their credit decline was supply-driven and nonlocal, increases the explanatory power of the credit supply instrument, while leaving coefficient estimates in the second-stage essentially unchanged.

In measuring the exposure of counties to lenders, I focus on large multimarket lenders operating in multiple counties who did not file for bankruptcy during the crisis. Specifically, I include lenders operating in at least 100 counties in 2007, and who issued over $1 billion in mortgage originations in the same year. Table 1.2 gives a summary of lender-level statistics. The 56 lenders account for 75 percent of mortgage lending over 2005-2007, so they cover the majority of lending by market share, even though there were over 6,000 mortgage lenders in that period. I roll up the remaining small institutions into a single entity.

1.3 Differences in Lender Supply

There were substantial differences in supply across lenders during the recession. Some lenders almost fully halted originations, while a few even expanded. For example, mortgage originations fell by 69 percent at Citibank but increased by 17 percent at US Bank (Table 1.2). The empirical challenge, a variant of the reflection problem in Manski (1993), is that those differences could reflect borrower characteristics rather than differences in lender supply. For example, it is possible that US Bank's typical customers experienced above-average credit demand during the recession. The main empirical strategy is to estimate lender fixed effects explaining variation in credit changes during the recession, while holding various characteristics of loans constant including the location of the property via locality fixed effects; other work employing similar methods includes Khwaja and Mian (2008), Greenstone, Mas, and Nguyen (2015), and Amiti and Weinstein (2018). This strategy exploits the richness in the HMDA data which provides originations at the locality-lender level and includes various loan characteristics.

The lender fixed effects reveal substantial differences in aggregate supply across lenders. They are largely driven by differences in lenders' funding strategy: reliance on funding sources that proved fragile in the crisis, such as wholesale debt and private loan sales in the secondary market, explain close to 75 percent of the variation in lender supply. In contrast, credit growth in the boom years (2003-2006) does not help explain either differences in lender supply or credit growth over 2007-2010, as shown in Figure 1.3. Therefore, I interpret the supply differences as largely reflecting exposure to unexpected funding cost shocks during the recession.

Specifically, I estimate versions of the following linear model that specifies credit changes during the recession as a function of lender fixed effects, locality fixed effects, and lender-locality interaction effects:[7]

$$\Delta L_{i,b} = \alpha_i + \phi_b + \gamma D_{i,b} + \upsilon_{i,b} \tag{1.1}$$

where $\Delta L_{i,b}$ are percent changes in mortgage credit originations at the county-lender level over 2007-2010; α_i are locality fixed effects (county or census-tract); ϕ_b are lender fixed effects; and $D_{i,b}$ are prerecession county-lender characteristics. The parameters of interest are those associated with the vector of lender fixed effects ϕ_b, which capture the idiosyncratic lender factor common across localities explaining variation in credit changes, net of locality fixed effects and prerecession county-lender characteristics.

The model captures many of the reasons for variation in credit changes at the lender-locality level during the Great Recession. For example, if originations to a locality declined sharply because of deteriorating local economic conditions – declines in local productivity, house prices, or credit scores – that will be captured by the locality fixed effects α_i. Similarly, if originations decline because it is difficult for lenders to fund new mortgages, that would be captured in the lender fixed effects ϕ_b. It is also possible that the variation is driven by interaction effects $D_{i,b}$ – for instance, Citibank's traditional borrowers could have tended to experience below-average credit demand shocks, even within localities.

In the baseline specification, I control only for county fixed effects. In this case the identifying assumption is that within-county credit demand shocks are uncorrelated with lender shocks. For example, supply contractions for Citibank

[7] The lender fixed effects are estimated using 30,161 county-lender observations, for the 56 lenders in the sample, and for county-lender pairs where the dollar value of originations is larger than $1 million. The lender fixed effects explain about a fifth of the variation in within-county lending changes over 2007-2010.

would be overestimated if their borrowers tend to be low-income, and low-income borrowers experienced worse credit demand shocks than average, even within-counties. To address this possibility, I estimate equation 1.1 using only high income loans, but estimates are very similar. Specifically, I estimate equation 1.1 using only loans to borrowers with income over $70,000 the median income of borrowers in 2007. The correlation coefficient between the lender fixed effects estimates in the baseline and the specification with only high income loans is 0.96; see Figure 1.4. When using only low-income loans (borrower income below $70,000), the correlation coefficient is also high, 0.92.

I also estimate equation 1.1 using census tract fixed effects rather than county fixed effects. Census tracts are statistical subdivisions of counties, each generally having a population size between 1,200 and 8,000 people. Census tracts are smaller and are more homogeneous than counties.[8] The lender fixed effects estimates when using census tract fixed effects are also highly correlated with the baseline (0.91). This shows that using a more detailed local control for changes in credit demand has very little bearing on the lender fixed effects estimates. Another alternative is to directly control for differences in the pre-recession profile of borrowers and lenders via county-lender characteristics $D_{i,b}$. The county-lender characteristics observed in HMDA are borrower income, fraction of loans classified as being high-risk, race, type of loan (owner-occupier), and credit growth in the peak boom years 2003-2006 by county-lender. When including $D_{i,b}$ in equation 1.1, the lender fixed effects estimates are again highly

[8] I rank census tracts within a county by borrower income, and divide the census tracts into four equal-sized groups by income, i.e. the top quartile consists of the high-income census tracts in the county. Census tract-income groups are more homogeneous than the county – in 2007, the median within-group standard deviation of HMDA borrowers in the census tract-income groups was $92 thousand, 27% lower than in counties.

correlated.

Table 1.4 shows sample statistics for a selected sample (by size) of 30 of the 56 lenders used in the paper. Column 2 provides percent changes in national mortgage originations over 2007-2010. Column 4 ranks lenders by the lender fixed effects estimates; above-average lender fixed effects indicate above-average supply. Changes in the ranking (going from Column 3 to 4) indicate differences in the degree to which national changes in mortgage originations were driven by geographic variation in exposure to credit demand shocks. For example, the drop in Bank of the West's ranking from 6^{th} to 15^{th} (from Column 1 to Column 2) indicates that lending changes for this bank remained relatively robust in the recession partly because of its exposure to above-average geographies (in this case the Midwest). Conversely, the improvement in the ranking of JPMorgan Chase from 50^{th} to 41^{th} indicates that part of its national decline in originations was driven by exposure to underperforming areas.

1.3.1 Funding Fragility and Differences in Supply

What explains the dispersion in aggregate supply across lenders, the variation in the lender fixed effects ϕ_b? In line with previous research documenting that low liquidity contributed to lower credit issuance during the crisis (Ivashina and Scharfstein 2010; Cornett et al. 2011; Brunnermeier 2008; Gorton and Metrick 2012; Kacperczyk and Schnabl 2010; Ramcharan, den Heuvel, and Verani 2013), this section shows mortgage lenders were more likely to cut supply during the recession if they relied on funding sources that proved fragile in the crisis. As discussed in Passmore, Sherlund, and Burgess (2005), mortgage loans can usually be funded in one of three ways: (i) via loan sales in the secondary market, or through balance sheet retention; (ii) if kept in the balance sheet, through wholesale debt or deposit-like liabilities; (iii) if sold in the secondary market, through loan sales to the GSEs (e.g. Fannie Mae, Freddie Mac, Ginnie Mae), or through sales to private buyers. I measure each of these three funding strategies by combining lender data from HMDA and the Federal Reserve's FRY-9C.

Table 1.5 reports results from regressions of differences in lender supply (ϕ_b) against differences in funding strategy over 2005-2007 (see also Figure 1.5) for the banks in the sample. Column 1 shows 72 percent of the variation in supply differences is explained by variation in reliance on wholesale debt, loan sales in general, and particularly sales to private investors.[9] Column 2 shows that lower prerecession capital ratios are also associated with declines in credit supply, though this factor is relatively minor, judging by its 5 percentage point contribution to the R-squared (Column 2). Column 3 shows that, in contrast, prerecession credit growth (over 2003-2006) is not helpful in explaining variation in differences in supply during the Great Recession. Observations are weighted by the dollar amount of mortgage originations in 2007, although the weighting is not critical, as the last column shows.

I measure bank-level exposure to wholesale funding as the ratio of non-core funding (sum of large time deposits, foreign deposits, repo sold, other borrowed money, subordinated debt, and federal funds purchased) to total assets, from the Federal Reserve's FRY-9C form, a standard definition in the literature (Irani and Meisenzahl 2015). To measure lender exposure to the secondary market, I use data from HMDA, which provides loan sales in the secondary market by

[9] In complementary work, Dagher and Kazimov (2012) find that mortgage lenders more reliant on wholesale funding were more likely to reject mortgage applications during the recession, after controlling for various borrower characteristics.

year and type of buyer. Reliance on loan sales is measured as the share of loans originated and sold to total originations over 2005-2007. Exposure to private securitization is measured as the ratio of private investor loan sales to total loan sales over 2005-2007.[10]

Measuring reliance on loan sales to private buyers is important since private-label residential mortgage securitization, which funded about 30% of mortgages over 2005-2007, went to essentially zero in 2008-2010 (Frame et al. 2015); see also Avery, Brevoort, and Canner (2011), and Nadauld and Sherlund (2009)). Because private investors stopped purchasing nongovernment-insured mortgages, lenders reliant on those sales likely cut supply during 2008-2010. For example, Calem, Covas, and Wu (2013) find that banks who were pre-recession more dependent on loan sales experienced more severe declines in jumbo lending, which are loans too large to be purchased by GSEs, and thus can only be sold to private investors, during the recession.

Loan sales to GSEs also became more expensive. G-fees, the monthly insurance fee GSEs charge as a fixed fraction of the loan balance, increased from about 20 basis points in 2005-2007 to 30 basis points in 2008-2010 (Fuster et al. 2013). Putback risk also increased in 2008. Lenders are required to repurchase loans sold to GSEs if it is found that those loans fail to satisfy original underwriting standards. While putbacks were rare, they rose during the recession, with Fannie Mae estimating that 3.7 percent of single-family loans purchased over 2005-2008 were putback to lenders, whereas the figure in other periods tended to be less than 0.5 percent[11]

[10] Private loan sales are defined as loan sales to any buyers excluding FNMA, FAMC, GNMA, FHLMC, and lender affiliates.

[11] source: Fannie Mae 10-K 2013, p. 143

1.4 The Nonlocal Lending Shock

Differences in lender supply affected counties differently, because of variation in the intensity of preexisting county-lender relationships, as measured by market shares prior to the recession. The credit supply instrument – the nonlocal lending shock – is the weighted average, for county i, of lender supply shocks in the recession ϕ_b (from equation 1.1). The weights are county-lender 2005-2007 mortgage origination market shares. The sum is taken over all large multimarket lenders in the sample B, as discussed in section 2.3:

$$Nonlocal\ Lending\ Shock_i = \sum_B Share_{i,b}\phi_b \tag{1.2}$$

Counties had below-average access to mortgage credit, all else equal, if they had existing relationships (as measured by 2005-2007 market shares) with lenders with below-average supply in the recession. New lender entry would work towards offsetting the decline in credit supply by the locality's traditional lenders. In the extreme case of perfectly substitutability, lender entry would fully offset the reduction in supply by the locality's traditional lenders.

The instrument, however, is not weak with the first stage F statistic in the baseline over 20. I provide evidence of both highly persistent county-lender relations prior to the crisis, and of limited new lender entry during the recession. First, county-lender market shares are highly persistent year-on-year. Table 1.6 shows results from regressing 2007 county-lender market shares on 2005 shares. Column 1 shows that 2005 shares explain 92 percent of the variation in 2007 shares, with the coefficient on the 2005 shares equal to 0.91. The left panel of Figure 1.6 plots 2007 shares against 2005 shares. Moreover, the relationship between 2005 and 2007 shares is highly stable across localities. The correlation coefficient and R-squared are very similar when focusing only on high credit score counties or only low credit score counties (Columns 2 and 3), or when using county fixed effects (Column 4). The persistence in credit relationship extends to at least the early 2000s. The right panel in Figure 1.6 plots 2000 market shares against 2007 market shares; there is a strong positive association, with 2000 shares explaining 71 percent of the variation in 2007 shares.

As for limited entry, I find few cases of lenders entering new counties in the recession: of all county-lender pairs in 2008-2010, only 7.85% were new matchings. The lack of entry suggests substantial switching costs across lenders during the recession. Part of the reason for low new entry may be that only a handful of lenders were expanding during the recession. Because most lenders were contracting, they may not have been looking to expand into new localities.[12] The contraction in lending by many mortgage lenders, particularly the larger ones, is also documented in Gete and Reher (2016) and Chen, Hanson, and Stein (2017).

The findings in this paper on persistent credit relationships and limited entry during the recession are in line with the literature documenting stickiness in mortgage credit relationships and limited shopping in the mortgage market in spite of significant price dispersion. In a survey of recent mortgage borrowers, Alexandrov and Koulayev (2017) report that close to half of the borrowers did not do any shopping. Woodward and Hall (2012) also find that borrowers engage in too little shopping, and "sacrifice at least $1,000 by shopping from too few brokers." Lacko and Pappalardo (2010) shows that mortgage borrowers

[12] These statistics are based on the 56 lenders in the sample as described in section 2.3.

are often severely uninformed about key costs associated with getting a mortgage, with half of respondents having problems identifying the loan amount, and two-thirds being unaware of prepayment penalties, for example. Moreover, Mondragon (2018) and Nguyen (2014) find evidence for stickiness in the mortgage market, in line with the large literature showing substantial switching costs for firms, as recently discussed in Chodorow-Reich (2014).

The main concern with instrument validity is that the credit supply instrument, county exposure to lender supply shocks, may be correlated with unobserved characteristics of counties affecting employment. It would be sufficient (but not necessary) if lender location is randomly distributed across counties. That is unlikely to be the case, however. Below-average suppliers in the recession may have been more likely to locate in subprime counties (for example) prior to the crisis. To the extent I can observe and control for the fraction of subprime borrowers (and other relevant local characteristics), I can isolate the 'as good as random' variation in lender location. To that end, I employ a detailed set of prerecession county characteristics, including the subprime share, that explains close to 60 percent of the cross-sectional variation in mortgage credit changes over 2007-2010, described in Table 1.3. The controls include: measures the run-up in house prices during the boom; industry composition; loan characteristics such as local incidence of FHA or investor loans; demographics; and measures of local lending competitiveness.[13] Figure 1.7 is a map of the nonlocal lending

[13] Previous literature has established that different household characteristics are associated with the severity of the housing boom and bust. For the incidence in subprime lending, see: Keys et al. (2010),Demyanyk and Hemert (2011), Dell'Ariccia, Igan, and Laeven (2012), Gerardi, Shapiro, and Willen (2008), and Mian and Sufi (2009). For the growth in household debt to income, see: Mian and Sufi (2014), and Carroll and Kimball (1996). For demographics: Elsby, Hobijn, and Sahin (2010). For loan characteristics: Haughwout et al. 2011; Chinco and Mayer 2016; Bhutta 2015; Bhutta and Ringo 2014

shock, after controlling for a detailed set of county covariates. The map appears balanced with no apparent trends by region.

Conditional on the detailed set of county observables used in the paper, I find evidence consistent with 'as good as random' county-lender matching both in the boom and before. The results in the paper are robust to measuring county exposure to lender shocks using 2000-2002 shares (instead of 2005-2007 shares). This addresses the concern that risky lenders may have located in risky counties during the housing boom. As for potential non-random county-lender matching before the 2000s, I produce results using no, region, division, or state fixed effects; substantially different estimates would be evidence of correlated fixed effects at regional levels for employment outcomes and lender location i.e., regional county-lender matching. However, estimates are consistent across specifications. As discussed shortly, I perform various other checks that find support for the exclusion restriction, including 'placebo' tests; controlling for declines in small business lending in the recession; and the inclusion of large, failed lenders (e.g. IndyMac) in the analysis.

1.5 Empirical Framework and Results

I now discuss results based on the following 2SLS specification:

$$\Delta Outcome_i^j = \theta X_i + \beta \Delta \widehat{Credit}_i + f_s + \epsilon_i \tag{1.3}$$

$$\Delta Credit_i = \delta X_i + \rho Nonlocal\ Lending\ Shock_i + f_s + v_i \tag{1.4}$$

where observations are at the county i level; changes are over 2007-2010 for different outcome variables j (house sales, house prices, employment) each estimated separately; and f_s are fixed effects that could be at the region, division, or state level – I report results for each. Table 1.3 defines the set of prerecession county controls X_i as well as the outcome variables. The nonlocal lending shock is the credit supply instrument defined in equation 1.2. All of the outcome variables are expressed as percent changes over 2007-2010. For employment categories and the home price index, changes are taken between 2007Q4 and 2010Q4. For mortgage credit (a flow) changes are taken between the average dollar flow over 2008-2010 with respect to the value in 2007.[14] Mortgage flows are deflated using the GDP deflator.[15]

Out of roughly 3,200 counties, I use data on slightly over the 1,000 largest counties in the U.S. (those having over 15,000 households in the 2000 Decennial Census), which account for 85% of aggregate employment. I drop states having 3 or fewer counties, to have at least a few observations per state for the specifications that use state fixed effects. Observations are weighted by the number of employed workers in 2006, though results are very similar without weighting.[16] Extreme observations (1% from each tail) are dropped from each dependent variable.[17] Standard errors are clustered at the division level to allow

[14]Using 2005-2007 as the base period produces nearly identical results, for example, the correlation between $\Delta Credit_i$ using 2007 as the base period and using 2005-2007 as the base is $\rho = .87$. Table 10 in the Online Appendix reports the main estimation results using 2005-2007 as the base.

[15]Alternatively, $\Delta Credit_i$ could be defined as the percent change in the number of mortgage originations, with very similar results; Table 11 in the Online Appendix shows the main 2SLS results when doing so.

[16]Table 12 in the Online Appendix reports unweighted results for counties with over 40,000 households in the 2000 Decennial Census – these close to 500 counties account for 76% of total employment.

[17]For example, I drop house price growth outliers from the house price regression, but I

for correlated shocks within broad geographic regions due to, for example, state or division-specific institutional arrangements and spatial correlation.[18] Estimates are robust to alternatives, such as clustering at the commuting zone level (Table 13 in the Online Appendix).

1.5.1 First Stage Results

The nonlocal lending shock has significant independent explanatory power over local changes in mortgage credit in the Great Recession, consistent with high switching costs across lenders. Table 1.7 reports first stage regression results; all the controls listed in Table 1.3 are included (e.g. the share of subprime borrowers, measures of the severity of the housing boom, and various demographic, industry, and loan characteristics) though only the nonlocal lending shock coefficient estimates are reported, to economize on space. Columns 1-4 include varying degrees of spatial fixed effects, ranging from none (Column 1) to region, division, and state fixed effects (Columns 2-4 respectively). The R-squared is reasonably high in all specifications (60 percent or higher), indicating that the regression controls are helpful in explaining variation in mortgage credit issuance. Across specifications, the coefficient estimate on the instrument is positive and strongly significant. For example, in the specification without spatial fixed effects (Column 1), a 10 percent reduction in the nonlocal lending shock is associated with a 4.99 percent decline in mortgage credit issuance; the

---

don't drop those counties from the private employment growth regression (unless they are also outliers in that variable). The only exception is growth in house sales for which I winsorize 5% of observations.

[18] The Census divides the US into 9 divisions – New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific.

first stage F statistic is slightly over 20.

The instrument has considerably explanatory power in all models with different spatial fixed effects. The first stage F statistic, associated with the hypothesis that the coefficient on the credit supply instrument is zero conditional on the observables, is over 10 in all the specifications, a rule of thumb commonly used to indicate weak instrument problems (Stock and Yogo 2002). The F statistic is lowest in the specification with state fixed effects (10.55), since this specification uses less information (only within-state variation in the instrument).

1.5.2 Effects of Supply Reductions on Residential Markets

Supply-driven, exogenous declines in mortgage credit are statistically associated with declines in home sales, home prices, and increases in delinquency rates as well as foreclosure rates. This is evidence of the negative effects of declines in mortgage supply on the health of local housing markets. The mechanism is that reductions in mortgage supply reduce the ability of households to buy homes and to refinance. Table 1.8 reports two stage least squares results for different housing market outcomes in the models with region fixed effects and standard errors clustered at the division level. I use region fixed effects in the baseline, though I provide results with other spatial fixed effects in the Online Appendix and in some cases throughout the paper.

Declines in credit supply are associated with declines in home permit issuance. Column 1 shows that a 10 percent reduction in mortgage credit (when instrumented using the nonlocal lending shock) is associated with a 10.37 percent decline in the issuance of new residential permits – essentially a one-to-one effect. This is evidence that households were unable to offset the reduction in credit availability originating from nonlocal sources by borrowing from private sources or from lenders other than their traditional, prerecession lenders.

Declines in mortgage credit are also associated with declines in home prices. A 10 percent reduction in mortgage credit is associated with a 5.92 percent decline in home prices.[19] The effect operates through the extensive margin – fewer loans were taken out, which led to lower housing demand and caused declines in home prices. Measuring changes in mortgage credit using declines in the number of loans, rather than in the real dollar value, yields nearly identical results.[20]

Delinquency rates and foreclosure rates also increased more in counties with below-average supply. Table 1.8 shows that a 10 percent decline in mortgage credit is associated with 1.33 and 0.79 percentage point increases in delinquency and foreclosure rates. This is evidence of the contractionary effects of reductions in mortgage supply on the health of local housing markets. The fall in home prices induced by the credit shock would make it more likely for households to go underwater.

In the Online Appendix I present results for the each dependent variable with no fixed effects, region, division, or state fixed effects. The main conclusions are essentially the same. The point estimates are very similar. For example, a 10 percent reduction in mortgage credit is associated with a 5.10, 5.92, 6.99, and

[19]This is consistent with other articles finding that supply-driven changes in credit have real effects on home prices, such as Favara and Imbs (2015), Mian and Sufi (2011), Adelino, Schoar, and Severino (2012), Favara and Imbs (2015), DiMaggio and Kermani (2014), Anenberg et al. (2016), Vojtech, Kay, and Driscoll (2016), Passmore and Sherlund (2016), and Kung (2015).

[20]The coefficient estimate in the model with region fixed effects when using declines in the real dollar value of mortgages is .592 while it is .601 when using declines in the number of loans.

5.59 percent decline in home prices in the models without, region, division, and state fixed effects respectively (Table 2 of the Online Appendix).

1.5.3 Effects of Supply Reductions on Employment

Declines in mortgage supply contributed to the job losses in the recession, though to a moderate extent. The job losses explained by the mortgage shock are concentrated in construction and financial services, a category of employment where over a third of workers are real estate intermediaries. The likely mechanism is that reductions in supply caused declines in housing demand, which negatively affected employment in industries reliant on housing demand. As evidence for this, I find that the construction losses are stronger in areas where housing supply is more elastic, that is, in areas where construction responds more to changes in housing demand. Overall, a supply-driven plausibly exogenous decline in mortgage credit issuance is associated with a 1 percent decline in total private employment. Using the in-sample variation of the nonlocal lending shock, I estimate that about 15 percent of the employment losses in the Great Recession can be explained by declines in mortgage supply.

Weak mortgage supply contributed to job losses in the construction sector. Table 1.9 shows that a 10 percent decline in mortgage credit originating from nonlocal sources is associated with a 5.14 percent decline in construction employment for the model with region fixed effects, with results being similar for the other specifications. The mechanism is that declines in mortgage supply reduce housing demand, which is associated with lower employment in construction.

The employment losses in construction were, for a given decline in instrumented credit, more severe in areas where housing supply is more elastic.[21] That is, in areas where construction responds more strongly to changes in housing demand, the employment effects of a given credit decline were stronger. To see this, I focus on the sample of counties for which the Saiz (2010) measure of the elasticity of housing supply is available.[22] Table 1.10 reports results for changes in home permits for new construction and construction employment for the model with region fixed effects. The coefficient estimate is positive for the interaction of credit changes and housing supply elasticity and significant at the 5% level for construction employment (Column 2). For the permits model (Column 4), the interaction is also positive, though marginally insignificant (p-value = .169). That is, the same relative decrease (increase) in credit is associated with lower (higher) permit issuance and construction employment in areas with higher housing supply elasticities. This is evidence for the mechanism that reductions in mortgage supply reduced housing demand and contributed to employment losses in construction.

Declines in mortgage supply also caused job losses in finance. Table 1.9 shows that a 10 percent reduction in mortgage credit is associated with a 3.90 percent decline in employment in financial services in the model with region fixed effects. The likely mechanism again is that reductions in supply negatively

[21] I add the interaction of credit changes and the housing supply elasticity to the regression model with region fixed effects. For the two endogenous regressors (credit changes and the interaction of credit changes and the housing supply elasticitiy), I use two instruments – the nonlocal lending shock, and the interaction of the nonlocal lending shock with the housing supply elasticity.

[22] Saiz (2010) estimates housing supply elasticity as a nonlinear combination of data on physical and regulatory building constraints and population levels in 2000 at the metro area level.

affected housing demand, and therefore demand for housing intermediaries.

Via the effects on construction and financial employment, declines in mortgage credit led to declines in total private employment. Table 1.9 shows two stage least squares results for different employment categories, including total private employment. Column 3 shows that a 10 percent reduction in mortgage credit originating from nonlocal sources is associated with a significant 1.14 percent decline in total private employment. The models with other types of spatial fixed effects have similar point estimates, as reported in Table 7 of the Online Appendix, though confidence intervals are wider, especially when division or state fixed effects are used. In the specification with state fixed effects, for example, a 10 percent reduction in instrumented mortgage credit is associated with a 0.73 percent decline in employment. Standard errors tend to be larger in the state fixed effects specifications, since they use less information (only within-state variation). The point estimates, however, tend to be very similar. The Online Appendix reports estimates for all of the dependent variables discussed in the paper for specifications with no spatial fixed effects, region, division, and state fixed effects.

Declines in mortgage supply are only weakly associated with declines in employment in other, broader employment categories – 'other employment' (total private excluding construction and finance) and nontradable employment, which mostly consists of local retail and food. These are shown in Table 1.9, Columns 4 and 5 respectively. The coefficient estimates are close to zero and not significant. That is, the negative shock on local real estate markets did not appear to significantly spillover to broader local employment categories. One possibility is that the real estate shock did have large spillover effects, but that those effects were nonlocal, and were instead dispersed through localities through the tradable sector. However, there is little evidence that the local real estate shock had large spillover effects on the local nontradable sector (Column 5). In Boldrin et al. (2016) the spillover between a housing shock to the rest of the economy depends on the elasticity of substitution between consumption and housing. The results in this paper suggest the (local) elasticity is relatively low.

The elasticity estimates of other and nontradable employment also contrast with their OLS counterparts, which are about three times larger and strongly significant, with t-statistics ranging from 3 to 8 across specifications, as shown in Table 1.11. That the OLS coefficients are larger suggests that they are biased upward, due to reverse causality – employment losses may lead to declines in mortgage issuance. The credit supply instrument is strong, and helps predict declines in real estate activity, such as declines in home permits, home prices, and construction employment. But it does not help explain substantial job losses in industries less directly related to real estate. This ameliorates concerns about reverse causality – if local employment shocks were correlated with the instrument, then the 2SLS estimates for broad employment categories would likely be large and significant.

In parallel work Mondragon (2018) also estimates the county level elasticity of employment with respect to mortgage supply during the recession. We both find that reductions in mortgage supply mattered for employment in the recession, though the estimated effects in Mondragon (2018) are substantially higher. He estimates that a 10 percent decline in instrumented mortgage credit is associated with a 3 percent decline in employment, an elasticity two times as large as the OLS counterpart, and three times as large as my own estimate.[23] The main difference between the papers is the credit supply instrument; his instrument is prerecession exposure to Wachovia Bank, a troubled lender acquired by Wells Fargo in late 2008.[24] One reason his estimates are likely larger is 'bad' bank in a 'bad' region matching – Wachovia had a larger presence in states in the South Atlantic such as Florida, South Carolina, and North Carolina where job losses were among the worst in the country.

The Wachovia instrument significantly weakens when controlling for characteristics of localities correlated with both Wachovia location and employment losses during the recession. For example, using only division or state fixed effects greatly diminishes the statistical power of the Wachovia instrument. To see this, I obtain Wachovia 2005-2006 purchase shares from HMDA and restrict the sample to counties in the South and East. The first stage F statistic associated with the Wachovia instrument is 14.47, absent other controls including regional fixed effects. When including division (state) fixed effects, the F statistic drops to 4.33 (0.93).[25] In contrast, the results in my paper are very similar when using no fixed effects, or region, division, or state fixed effects. Moreover, it is not the case that the results are different because Wachovia was a particularly troubled lender. In fact, Wachovia was acquired by Wells Fargo, the strongest lender of the top 4. As discussed shortly, the results in this paper are very similar even

[23] These estimates replicate the earlier Mondragon (2014). In more recent versions, changes in Mondragon's specifications such as variable standardization and sample restrictions make replicating Mondragon (2018) less straightforward.

[24] Mondragon (2018) continues to use Wachovia as the key source of identification as in earlier versions (Mondragon 2014), though the more recent version uses a few other regional lenders as a robustness check.

[25] Observations are weighted by population in 2006, and standard errors are clustered by state. These results are available upon request.

when including large, failed lenders in the analysis such as IndyMac, which was not rescued by another institution.

1.5.4 Aggregate Implications

Overall, I find that reductions in mortgage supply could explain close to close to 15 percent of the employment losses in the U.S. over 2007-2010, or about 1.1 million of the jobs lost. This is evidence that reductions in mortgage supply mattered for employment. The imputation is based on a partial equilibrium aggregation exercise that answers the counterfactual question: what if counties, all else equal, had experienced the best credit shock in the sample – specifically, the credit shock of the counties in the top 5 percent of the distribution? The improvement in supply generates employment gains via the estimated elasticity of employment with respect to mortgage supply. This approach addresses the challenge that the level effect of supply reductions cannot be recovered from the cross-section by assuming that the top percentile of counties by the credit supply instrument represent a 'no credit shock' scenario. This is a standard aggregation exercise in this literature, with similar approaches in Chodorow-Reich (2014) and Mian and Sufi (2014). The estimate would be biased downwards if the top percentile counties also experienced a reduction in supply. The severe disruptions in mortgage supply in the recession affecting wholesale funding markets and loan sales in the secondary market suggest the assumption is conservative.

First, define the counterfactual employment change in county i, ΔEmp_i^{cf}, as the predicted employment if county i had experienced the nonlocal lending shock of county zero (NLS_0), rather than its own (NLS_i), after conditioning on all other observables X_i:

$$\begin{aligned}\Delta Emp_i^{cf} &= E[\Delta Emp_i | NLS_i = NLS_0, X_i] \\ &= \Delta \widehat{Emp}_i + \beta(\Delta \widehat{Credit}_i(NLS_0) - \Delta \widehat{Credit}_i(NLS_i)) \\ &= \Delta \widehat{Emp}_i + \beta\rho(NLS_0 - NLS_i)\end{aligned}$$

where $\Delta \widehat{Emp}_i$ denotes the fitted value from the private employment regression model with region fixed effects, β is the estimated elasticity of employment with respect to mortgage supply, and ρ is the coefficient on the nonlocal lending shock in the first stage regression. I then recover the end-period levels of employment corresponding to both the counterfactual and fitted changes in employment, using the initial-period employment level: $Emp_{i,2010Q4}^{cf} = Emp_{i,2007Q4}(1 + \Delta Emp_i^{cf})$ and $\widehat{Emp}_{i,2010Q4} = Emp_{i,2007Q4}(1 + \Delta \widehat{Emp}_i)$. Then, the total job loss explained by variation in the nonlocal lending shock is given by:

$$Total\ jobs\ lost\ explained\ by\ lending\ shock = \sum_{i:NLS_i < NLS_0} [Emp_{i,2010Q4}^{cf} - \widehat{Emp}_{i,2010Q4}] \tag{1.5}$$

The fraction of jobs lost that is explained by the lending shock is given by:

$$\frac{\sum_{i:NLS_i < NLS_0} [Emp_{i,2010Q4}^{cf} - \widehat{Emp}_{i,2010Q4}]}{\sum_{i:NLS_i < NLS_0} [Emp_{i,2010Q4} - Emp_{i,2007Q4}]} \tag{1.6}$$

The exercise indicates that the decline in mortgage supply can explain 14 percent of the employment losses in the Great Recession, when defining county zero as the 95th percentile county by the credit supply instrument, and using the coefficient point estimate $\beta = 0.114$ from the region fixed effects model. There is uncertainty around β, however. For example, $\beta = 0.073$ in the model with state fixed effects. Using the latter, the aggregation exercise suggests the mortgage credit supply shock explains 9 percent of the job losses during the recession. Alternatively, the 95 percent confidence interval for β in the region fixed effects specification ranges from 0.048 to 0.179; using this range, the decline in mortgage supply explains between 6 and 22 percent of the job losses during the recession.

Another important parameter choice is which counties are used as the 'no credit shock' reference. The baseline uses the 95th percentile as the baseline. If localities in the top 5 percent of the credit supply distribution also experienced a reduction in credit supply, the aggregation exercise will deliver an underestimate. When using the top 1 percent as a reference instead, the aggregation exercise suggests declines in mortgage supply can explain 21 percent of the job losses in the recession.

The bottom line of these aggregation exercises is that the reduction in mortgage supply likely aggravated the job loss during the recession, though moderately so. 14 percent of the total job losses is sizable – about 1 million jobs lost is hardly small – but it is far from the bulk of the job losses, as argued by Mondragon (2018) and particularly Mondragon (2014) which attributed about 60 percent of the total job losses (at a minimum) to household credit supply shocks. In sum, the evidence in this paper adds nuance to the debate of "what explains the job losses during the recession?" The answer provided by this paper is that mortgage supply shocks mattered, though moderately. This suggests that other factors – the decline in household net worth (Mian and Sufi 2014), increase in uncertainty (Baker, Bloom, and Davis 2015), or credit supply reductions to firms (Chodorow-Reich 2014) – likely explain the bulk of the job losses in the recession.

1.5.5 Robustness

I test for the validity and interpretation of the main results of the paper along several dimensions. As discussed, a concern is that lenders with below-average supply systematically located in counties with below-average employment shocks during the boom – perhaps risky lenders moved to risky counties during the boom years. I measure the credit supply instrument as in equation 1.2, with the same lender shocks during the recession ϕ_b, but this time using 2000-2002 market shares (instead of 2005-2007 as in the baseline).

$$Nonlocal\ Lending\ Shock_i^{2000-2002\ shares} = \sum_B Share_{i,b}^{2000-2002} \phi_b \tag{1.7}$$

Figure 1.8 plots the baseline credit supply instrument measure against the instrument measured with 2000-2002 shares; the R-squared is close to 64 percent. Table 1.12 reports 2SLS results based on county exposure to lender shocks, with the exposure measured in 2000-2002. For identification, the important thing is the point estimates are very similar, which is evidence that β^j are estimated consistently for different models j. The point estimates are indeed similar; for example, β is .106 in the total employment model (Column 3) while it is 0.114 in the baseline reported in Table 1.9, well within one standard error. The estimates are noisier – in the baseline, the first stage F statistic was 27.16 whereas in this specification it is 15.91 – as is expected, due to the noise in measuring lender location in the early 2000s rather than immediately prior to the crisis.

I also run 'placebo' tests on the first and second stage equations. First, I regress yearly changes between mortgage credit (2000-2013) at the county-level on the nonlocal lending shock and all the county controls in the baseline case. Figure 1.9 plots the coefficient estimates and associated 95 percent confidence intervals on a year-by-year basis. The mortgage credit shock helps explain credit changes over 2007-2008 and 2008-2009 only, and not during any of the prerecession years.[26] Second, I repeat the main 2SLS elasticity estimates, but now using left-hand side variables (e.g. employment changes) measured over the last two recessions: i) 1990-1992, during which the unemployment rate increased from 5.6 to 7.5 percent; and ii) 2000-2003, during which the unemployment rate increased from 4 percent (lowest since 1970) to 6 percent in 2003 – the previous two recessions also considered in Duygan-Bump, Levkov, and Montoriol-Garriga (2015). Table 1.13 reports elasticity estimates for construction and total private employment. If coefficient estimates are positive and significant, that would indicate counties with below-average supply during the Great Recession tend to experience below-average employment shocks during other recessions, possibly for other unobserved characteristics of localities. However, the estimates are insignificant, except for changes in total private employment over 1990-1992, though in this case the coefficient estimate has the opposite sign (negative rather than positive).

In the baseline results of the paper, I did not include institutions that filed for bankruptcy (and were not acquired by another lender), because the portion of lending changes that is nonlocal cannot be plausibly isolated for these lenders, since lending for these institutions fell by 100% everywhere (there is no variation

[26] Making a similar coefficient plot using a different dependent variable (e.g. total private employment) yields non-significant coefficient estimates. The reason is there are efficiency gains with lumping the recession years into a single cross-section. Results available upon request.

across localities). This is a conservative choice. The inclusion of these lenders might lead to biased elasticity estimates. On the other hand, their exclusion likely decreases the statistical power of the estimation approach. I add to the sample the ten largest multimarket lenders who failed over 2005-2010.[27] Table 1.14 reports two stage least squares estimates when the credit supply instrument includes these large failed lenders. Their addition leads to a small increase in the first stage F statistic. Moreover, the second stage estimates are very similar to the baseline. Some are a bit higher and some a bit smaller, though all within one standard error of the baseline estimates.

I also check whether coefficient estimates are statistically different when adding additional controls. In particular, I add squared and cubed terms of some of the most important drivers of the housing boom and bust identified in the literature: the runup in home prices over 2003-2006, 2006 debt-to-income, and the fraction of borrowers in a county with FICO scores less than 620. Table 1.15 reports the main regressions of the paper (with region fixed effects), this time including as additional explanatory variables the squared and cubed prerecession terms of these three variables. The results are essentially identical, ameliorating concern about omitted variable bias. The total private employment coefficient estimate is 0.126, compared with 0.114 in the baseline.

Finally, I show that the results in the paper are robust to controlling for realized declines in small business lending over 2007-2010, which I obtain from the Community Reinvestment Act dataset. I average the flow of new business originations over 2008-2010, and compute percent changes with respect to

[27] American Home Mortgage, New Century Financial, IndyMac, Fremont Investment, WMC Mortgage, Lehman, Ameriquest, Option One, First Magnus, and Taylor, Bean, & Whitaker Mortgage.

2007. Table 1.16 shows that controlling for the change in small business lending does not affect the main results of the paper. This is evidence that the mortgage credit shock discussed in this paper is carefully identified, and pertains specifically to changes in the availability of mortgage credit. The total private employment coefficient estimate is 0.111, compared with 0.114 in the baseline. Part of the reason why the two channels are distinct is that the exposure of localities to small business and mortgage lenders is only weakly correlated. In other words, the small business lenders to a locality are often not the same as the mortgage lenders. Figure 1.10 plots HMDA shares against CRA shares for the top 4 banks; they are only weakly correlated.

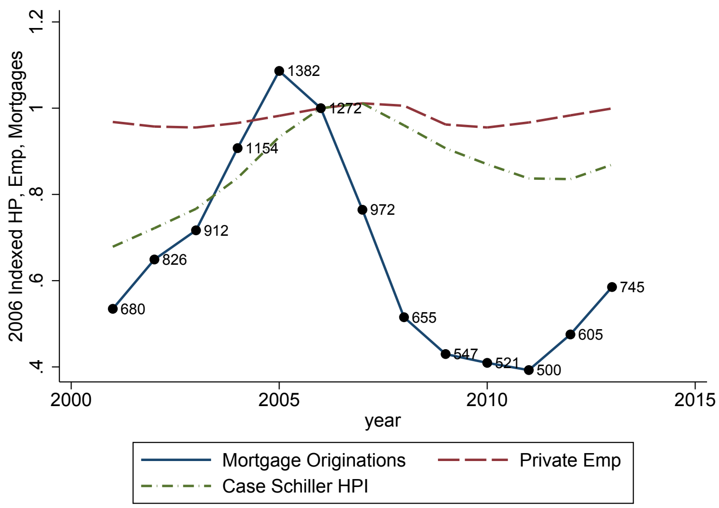
1.6 Conclusion

One of the leading narratives of the Great Recession is the credit crunch view – disruptions in financial markets limited the supply of new credit, which reduced the spending capacity of households and firms and lowered aggregate demand and employment, as discussed in prominent models of the Great Recession (Eggertsson and Krugman 2012; Guerrieri and Lorenzoni 2011; Midrigan and Philippon 2016). This paper contributes to this literature by empirically quantifying the employment effects of changes in mortgage credit supply. The emphasis on mortgages complements existing research the majority of which focuses on corporate credit supply shocks (Chodorow-Reich 2014; Greenstone, Mas, and Nguyen 2015).

To do so, I construct a county level mortgage credit supply instrument, which exploits two sources of heterogeneity: differences in the extent to which lenders cut supply in the Great Recession for nonlocal reasons, and variation in the intensity of county-lender relations coming into the recession. I then estimate the effect of changes in mortgage supply on employment, net of other possibly confounding factors affecting spending during the recession.

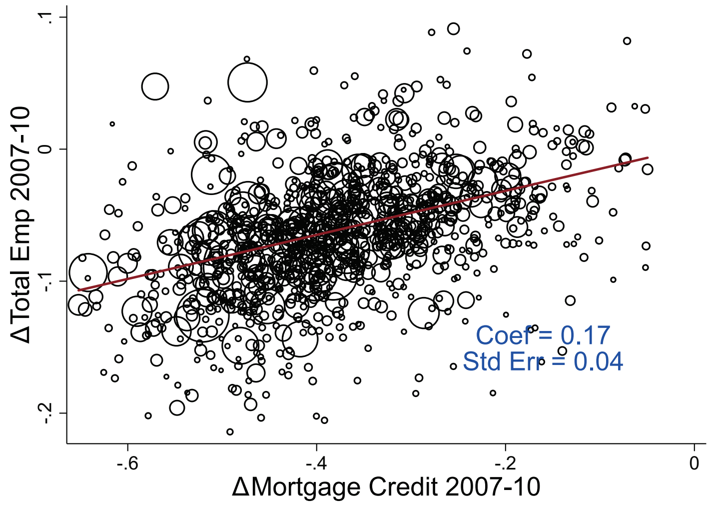
By quantifying the effects of mortgage supply reductions, this paper adds nuance to the debate on the drivers of the job losses during the Great Recession. Overall, the bottom line is that mortgage supply shocks mattered for employment, though only moderately so. Declines in mortgage supply caused declines in local real estate activity – in residential permits, house prices, and construction employment, for example – but the evidence does not suggest there were large spillover effects in other, broader employment categories. A partial equilibrium aggregation exercise, based on the estimated local elasticity of total private employment with respect to mortgage supply, indicates that the reduction in mortgage supply could explain close to 15 percent of the employment losses in the Great Recession, or about 1.1 million of the jobs lost. In other words, the reduction in mortgage supply likely aggravated the job losses to a meaningful extent. But, other factors – the decline in household net worth, increase in uncertainty, or credit supply reductions to firms – likely explain the bulk of the job losses in the recession, particularly in sectors less directly linked to real estate.

Figure 1.1: National Trends in Employment, House Prices, and Mortgage Originations



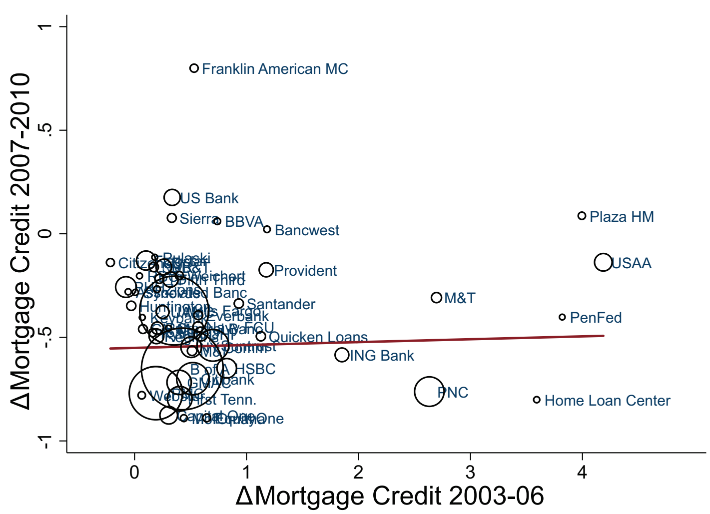
Mortgage originations are defined as the dollar value (in trillions) of originations for 1-4 residential loans for home purchase and improvement. Source: HMDA.

Figure 1.2: County Level Changes in Employment against Changes in Mortgage Credit Issuance, 2007-2010



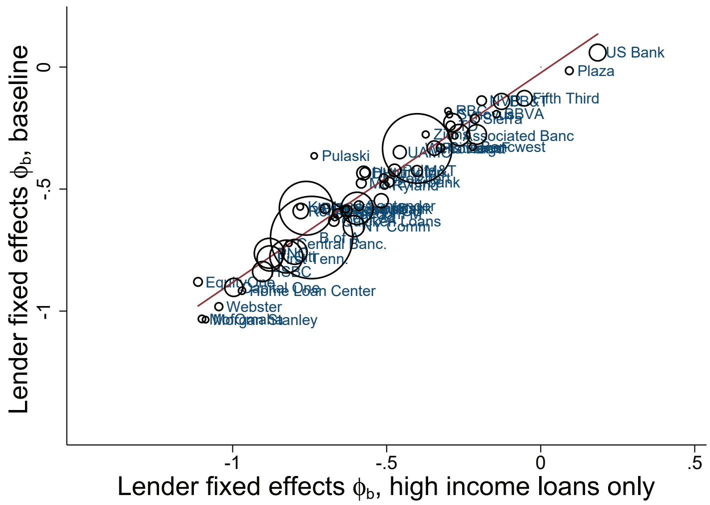
The figure plots changes in total private employment (y-axis) against changes in mortgage credit issuance (x-axis) over 2007-2010 at the county level for locations with over 15,000 housing units in the 2000 Census. The figure shows the linear coefficient estimate when regressing changes in employment on changes in mortgage credit issuance. Observations weighted by population in 2006. Standard errors clustered at the division level.

Figure 1.3: Credit Changes 2007-2010 Versus Changes in 2003-2006



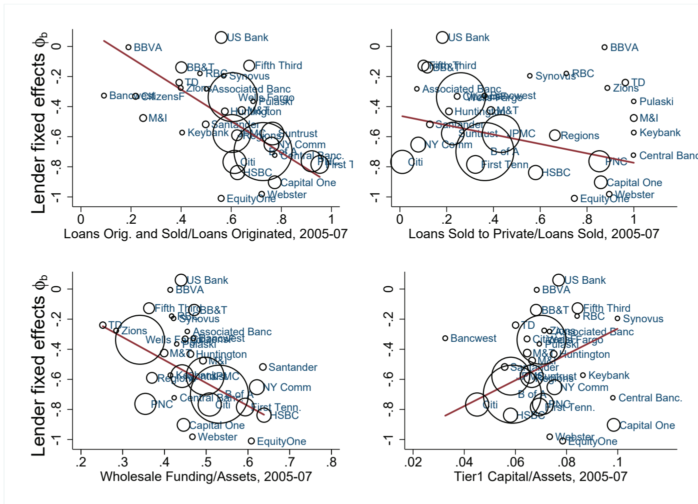
The figure plots changes in mortgage credit over 2007-2010 versus changes in mortgage credit over 2003-2006 for the large multimarket lenders in the sample.

Figure 1.4: Lender Fixed Effects Estimates in Baseline vs Only High Income Loan Specification



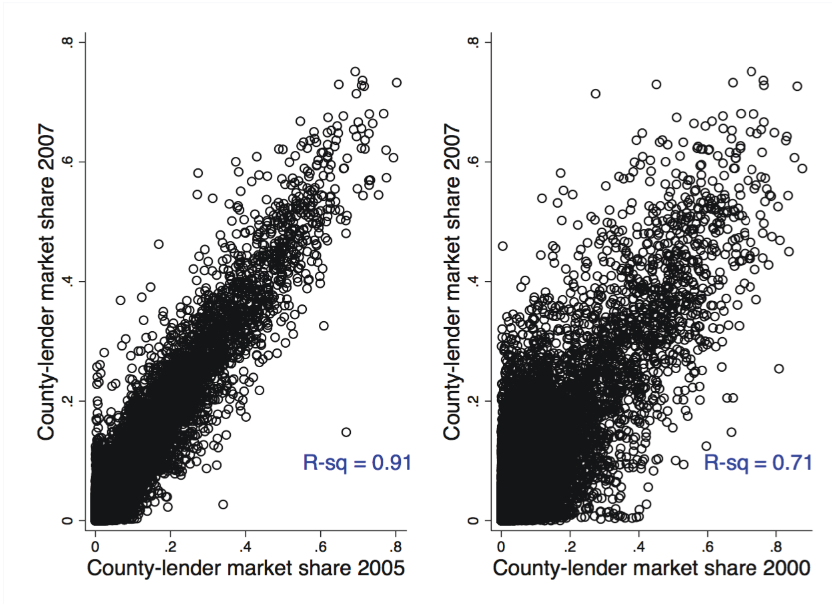
The figure plots lender fixed effects (equation 1.1) in the baseline (y-axis) against a specification that uses only high-income loans to estimate equation 1.1. Observations are weighted by the 2007 dollar value of mortgage originations.

Figure 1.5: Funding Fragility and Lender Supply



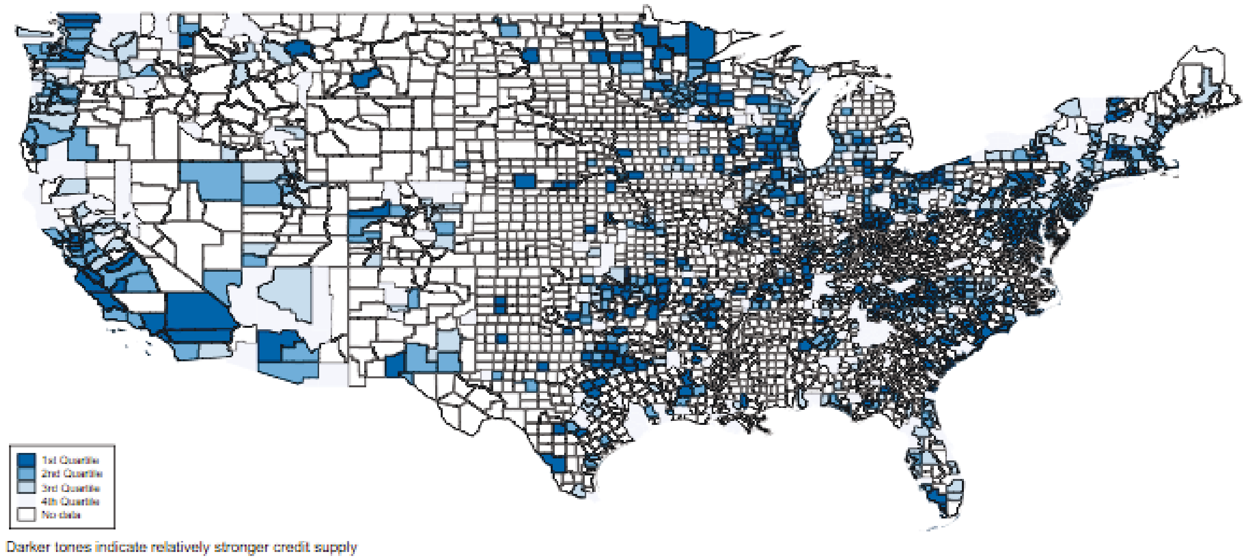
The variable on the y-axis measures differences in lender supply over 2007-2010, ϕ_b from equation 1.1. Variables on the x-axis are different measures of funding fragility over 2005-2007: ratio of mortgages originated and sold to total mortgages originated (top left); loans sold to private investors to total sales (top right); wholesale funding to assets (bottom left); and Tier 1 capital to assets (bottom right). Observations weighted by mortgage originations in 2007. The banks in the sample are large multimarket lenders located in at least 100 counties and with originations in excess of $1 billion in 2007.

Figure 1.6: Persistent Market Shares



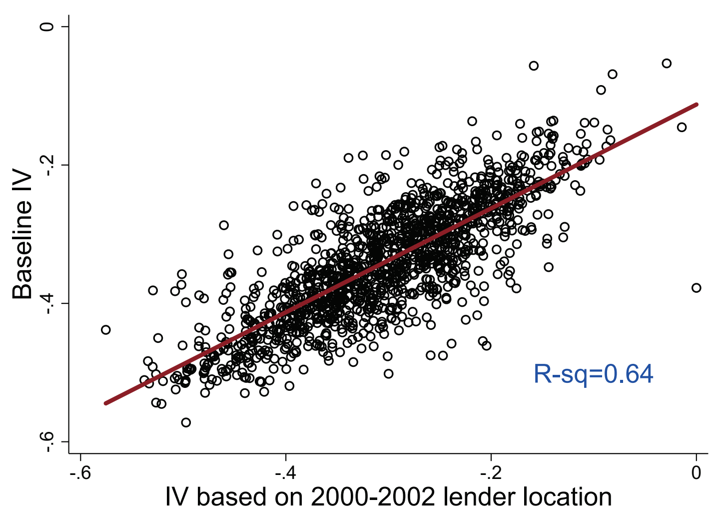
The left panel plots county-lender HMDA market shares in 2007 (y-axis) against market shares in 2005 (x-axis). The right panel plots county-lender HMDA market shares in 2007 (y-axis) against market shares in 2000 (x-axis). Lenders in the sample were located in at least 100 counties, issued over \$1 billion in mortgage originations in 2007, and did not fail during the crisis. Counties in the sample had over 15,000 housing units in the 2000 Census.

Figure 1.7: Nonlocal Lending Shock



The map plots the residual variation in the credit supply instrument (the nonlocal lending shock) after regressing the credit supply instrument on the county controls used throughout the paper and defined in Table 1.3. The instrument is defined in equation 1.2. The map sorts the nonlocal lending shock into quartiles for counties in the sample. Darker tones indicate relatively stronger supply. Missing observations left blank (in white).

Figure 1.8: Nonlocal Lending Shock using 2000-2002 Market Shares



This figure plots the baseline credit supply instrument (nonlocal lending shock) on the y-axis, against the credit supply instrument which measures lender location over 2000-2002 in the x-axis. The baseline instrument measures lender location using 2005-2007 county-lender market shares as defined in equation 1.2.

Figure 1.9: Regressing Yearly Mortgage Credit Changes on Nonlocal Lending Shock

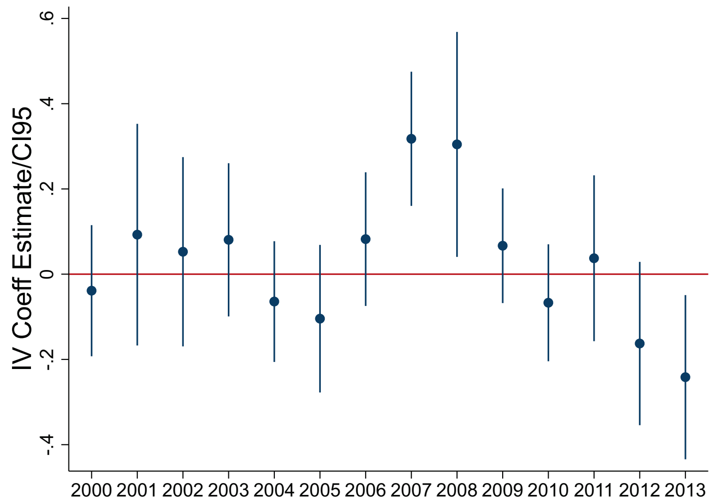
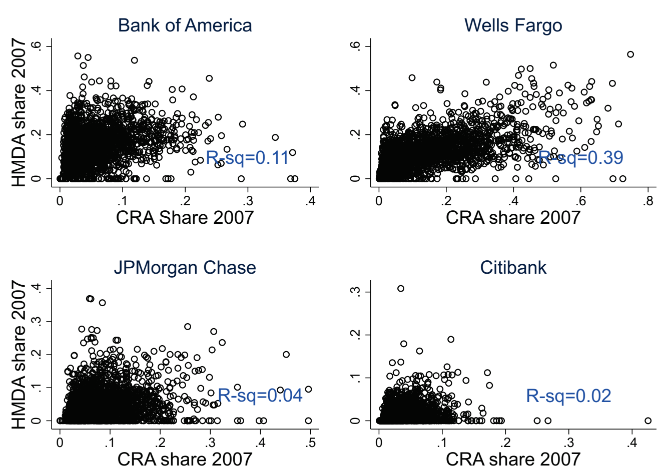


Figure shows coefficient estimates (ρ_t) and 95 percent confidence intervals when regressing yearly mortgage credit changes at the county-level on the nonlocal lending shock NLS_i and the other controls used in the baseline specification: $\Delta Credit_{i,t} = \rho_t NLS_i + \gamma X_i + \upsilon_i$ for $t = 2001, 2002, ...2014$

Figure 1.10: County-Lender Market Shares in HMDA and CRA



The figure plots mortgage 2007 market shares from HMDA (y-axis) against 2007 small business loan market shares from the CRA (x-axis) for each of the big-4 lenders.

Table 1.1: County Summary Statistics

	Mean	SD	p10	Median	p90	N
Dependent Variables, 2007-2010 percent changes						
Δ Private Emp	-0.065	0.048	-0.126	-0.063	-0.008	1063
Δ Construction Emp	-0.241	0.143	-0.415	-0.248	-0.058	1063
Δ Finance Emp	-0.080	0.089	-0.188	-0.082	0.024	1063
Δ Other Emp	-0.049	0.050	-0.111	-0.049	0.012	1063
Δ Nontradable Emp	-0.045	0.065	-0.119	-0.051	0.037	1063
Δ Home Prices	-0.139	0.103	-0.285	-0.126	-0.012	1063
Δ Delinquency +90 Days	0.039	0.023	0.018	0.032	0.070	1063
Δ Foreclosures	0.015	0.012	0.006	0.011	0.025	1063
Δ Mortgage Credit	-0.376	0.116	-0.516	-0.387	-0.218	1063
Prerecession Characteristics, 2006 levels and 2003-2006 percent changes						
Δ Home Prices	0.249	0.170	0.068	0.203	0.505	1063
Debt to Income	1.778	0.595	1.171	1.640	2.597	1063
Median Income (thousands)	\$48	\$12	\$37	\$45	\$66	1063
Median FICO	710	32	660	716	747	1063
% FICO <620	0.269	0.082	0.172	0.257	0.389	1063
% Owner-Occupied Loans	0.849	0.078	0.755	0.869	0.921	1063
% Conventional Loans	0.82	0.11	0.67	0.83	0.95	1063
% GSE-securitized Loans	0.66	0.14	0.49	0.69	0.78	1063
Construction Share of Emp	0.12	0.04	0.08	0.11	0.18	1063
Tradable Share of Emp	0.13	0.08	0.04	0.12	0.23	1063
% White Population	0.86	0.13	0.69	0.90	0.98	1063
% Educ $\geqslant$ College	0.15	0.05	0.08	0.13	0.22	1063

The table provides summary statistics for localities with over 15,000 households in the 2000 Decennial Census. For prerecession characteristics, level variables are measured in 2006 while percent changes are taken over 2003-2006 with 2003 as the base year. The change in delinquency and foreclosure rates is in percentage points. For stocks, changes are taken between 2010Q4 and 2007Q4. For flow variables, changes are taken between the average flow over 2008-2010 and the value in 2007.

Table 1.2: Lender Summary Statistics

	Mean	SD	p10	Median	p90	N
Δ Mortgage Credit 2007-2010	-0.38	0.30	-0.79	-0.37	0.01	56
#Counties 2007	487	475	121	279	1,117	56
Mortgage Credit 2007 (billions)	$12.12	$30.76	$1.08	$2.48	$24.39	56
Loan Sales/Loans Originated 2005-2007	0.68	0.25	0.36	0.70	0.99	56
Loans Sold to Private Investors/Loans Sold 2005-2007	0.60	0.37	0.08	0.64	1.00	56
Wholesale Funding/ Assets 2005-2007	0.44	0.11	0.34	0.44	0.61	31
Tier 1 Capital/Assets 2005-2007	0.07	0.01	0.06	0.07	0.09	31

This table provides summary statistics for the lenders in the sample, which are large multi-market lenders located in at least 100 counties and with originations in excess of $1 billion in 2007.

Table 1.3: Data Definitions

Variable	*Definition*	*Source*
Dependent Variables, 2007-2010 percent changes		
Mortgage Credit	By county-year, the dollar amount of originations for 1-4 residential loans for home purchase and improvement.	HMDA
Δ Credit	Percent change in average mortgage credit over 2008-2010 with respect to 2005-2007	HMDA
Δ Residential Permits	Percent change in average permits over 2008-2010 with respect to 2005-2007	Census
Δ House Prices	Percent change in house prices from 2007Q4 to 2010Q4.	CoreLogic HPI
Δ Emp^j	Percent change in employment category j from 2007Q4 to 2010Q4	QCEW
Prerecession Characteristics, 2006 levels and 2003-2006 percent changes		
Household Income	Median	IRS
FICO score	Median	Equifax CCP
Subprime	Fraction of households in a county with FICO score less than 620)	Equifax CCP
Household Debt-to-Income	Median household debt-to-income	Mian and Sufi (2014) Web Appendix
White population	Fraction of population identified as white	Census
College population	Fraction of population with a college degree or more	Census
Nonconventional Loans	One minus the fraction of loans issued over 2003-2006 identified as conventional loans	McDash
GSE-securitized Loans	Fraction of loans issued over 2003-2006 insured by GNMA, FNMA, or FHLMC	McDash
Owner-Occupied Loans	Fraction of mortgages over 2003-2006 identified as owner-occupied	HMDA
Herfindahl Index	Sum of squared market shares across lenders in county	HMDA
Δ # Lenders	Growth in the number of lenders per county over 2003-2006	HMDA
Δ House Prices	Growth in house prices over 2003Q4-2006Q4	CoreLogic HPI
Construction	Construction share of employment	QCEW
Tradable	Tradable share of employment, where tradable employment is defined as in Mian and Sufi (2014)	CBP
Unemp Rate	Unemployment Rate	BLS LAU
Level Home Prices	Log level median house price	Census
Level Employment	Log level of employed workers	QCEW
Level Mortgage Credit	Log level of mortgage originations	HMDA

This table provides definitions and sources for the data used throughout the paper. Unless otherwise specified, prerecession level variables are measured in 2006 while growth rates are taken over 2003-2006. Outcome variables are in percent changes over 2008-2010 with respect to a prerecession period. For stocks, changes are taken between 2010Q4 and 2007Q4. For flow variables, that change is taken between the average flow over 2008-2010 and the value over 2005-2007.

Table 1.4: Lender Rankings by Percent Changes in Mortgage Originations and Lender Fixed Effects Estimates

Lender	ΔOriginations, 2007-2010	Lender Ranking by ΔOriginations	Lender Ranking by Lender Fixed Effects Estimates	Originations, $billions 2007
US Bank	17%	2	2	7,449
Bank of the West	1%	6	15	1,260
Flagstar	-13%	7	13	10,368
Pulaski Mtg Co.	-15%	9	27	1,029
Citizens Financial	-15%	10	20	1,876
BB&T	-15%	11	7	6,836
Provident Funding	-18%	13	23	5,644
Toronto Dominion	-21%	16	17	2,211
Fifth Third	-21%	17	4	6,412
M&T	-34%	24	29	3,093
Wells Fargo	-37%	27	24	129,800
Everbank	-37%	28	18	2,307
UAMC	-38%	29	19	4,522
Navy FCU	-43%	30	33	3,300
Quicken Loans	-45%	33	34	2,838
Pulte	-45%	34	31	4,047
Freedom Mtg	-47%	36	37	2,463
DHI	-48%	38	28	5,086
Regions	-49%	39	44	6,305
Suntrust	-53%	40	39	27,855
NY Community	-54%	41	43	12,011
M&I	-55%	42	21	2,546
HSBC	-65%	44	52	10,888
Bank of America	-65%	45	45	182,100
Citibank	-69%	46	47	29,109
Ally Financial	-71%	47	46	16,627
PNC	-76%	49	49	24,105
JPMorgan Chase	-76%	50	41	77,644
First Tennessee	-79%	52	51	17,049
Capital One	-86%	53	53	8,817
Median across Lenders	-37%			$2,477
Standard Deviation	30%			$30,656

The table shows summary statistics for large nonfailed multimarket lenders in the sample (see Section 2.3). Column 2 ranks lenders by decline in new mortgage lending. Column 3 ranks lenders by differences in lender supply, ϕ_b from equation 1.1.

Table 1.5: Funding Fragility and Differences in Lender Supply

	Dependent variable: ϕ^b			
	(1)	(2)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Wholesale Debt/Assets 2005-2007	-0.356***	-0.267***	-0.220**	-0.424***
	(0.071)	(0.074)	(0.084)	(0.124)
Loan Sales/Originations 2005-2007	-0.491***	-0.537***	-0.646***	-0.662***
	(0.123)	(0.114)	(0.146)	(0.134)
Private Loan Sales/Originations 2005-2007	-0.380***	-0.452***	-0.528***	-0.559***
	(0.113)	(0.108)	(0.125)	(0.128)
Tier1 Capital 2005-2007		0.276**	0.330**	0.336**
		(0.113)	(0.121)	(0.145)
Δ Mortgage Credit 2003-2006			0.126	0.061
			(0.108)	(0.127)
Weighted	Yes	Yes	Yes	No
N	31	31	31	31
R-squared	0.72	0.77	0.79	0.68
Adj R-squared	0.69	0.74	0.74	0.62

The dependent variable measures differences in lender supply over 2007-2010, ϕ_b from equation 1.1. The explanatory variables measure the extent to which banks relied on fragile funding sources over 2005-2007, and credit growth over 2003-2006. Standard errors are in parentheses. Banks in the sample are large multimarket lenders as described in Section 2.3. All variables are standardized. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.6: Mortgage Market Shares are Highly Persistent Year-on-Year

	Dependent variable: County-Lender Market Shares 2007			
		Bottom FICO quartile	*Top FICO quartile*	*County FE*
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
2005 Market Shares	0.906***	0.909***	0.909***	0.905***
	(0.00)	(0.01)	(0.01)	(0.00)
County FE	No	No	No	Yes
R-squared	0.92	0.89	0.93	0.92
Observations	35651	8679	8577	35651

This table show results from regressing 2007 county-lender market shares on 2005 county-lender market shares. Column 2 restricts the sample to the low FICO score quartile, Column 3 to the high FICO score quartile, and Column 4 includes county fixed effects. The lenders in the sample are large multimarket lenders located in at least 100 counties and with originations in excess of $1 billion in 2007. Counties in the sample had over 15,000 households in the 2000 Decennial Census. Standard errors clustered at the county level. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.7: First Stage Results

	Dependent variable: Δ Mortgage Credit 2007-2010			
	No FE	*Region FE*	*Division FE*	*State FE*
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Nonlocal Lending Shock	0.499***	0.555***	0.437***	0.282**
	(0.11)	(0.11)	(0.12)	(0.09)
All other controls	Yes	Yes	Yes	Yes
R-squared	0.57	0.63	0.67	0.76
Adj. R-squared	0.56	0.62	0.66	0.75
First stage F stat	20.85	25.86	13.90	10.55
Observations	1045	1045	1045	1045

This table shows the effects of changes in mortgage credit, when instrumented using the nonlocal lending shock, on changes in local outcomes for the largest 850 U.S. counties. The nonlocal lending shock measures the exposure of counties to lender shocks (see text for details). Observations weighted by the number of employed workers in 2006. The dependent variable is winsorized 1 percent in each tail. Standard errors clustered by state. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.8: Housing Elasticities with respect to Mortgage Supply

	Dependent variables 2007-2010:			
	Δ Permits Coef./SE	Δ Home Price Coef./SE	Δ Delinq. Rate Coef./SE	Δ Foreclosure Rate Coef./SE
Δ Mortgage Credit 2007-2010	1.037*** (0.15)	0.592*** (0.12)	-0.133*** (0.04)	-0.079*** (0.02)
All other controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
R-squared	0.48	0.78	0.79	0.59
First stage F stat	28.23	26.03	27.44	29.32
Observations	951	1028	1031	1032

This table shows the effects of changes in mortgage credit, when instrumented using the nonlocal lending shock, on changes in local outcomes for the largest 850 U.S. counties. The nonlocal lending shock measures the exposure of counties to lender shocks (see text for details). All regressions include region fixed effects and all other observed characteristics of localities used in the other tables in the paper. Observations weighted by the number of employed workers in 2006. The dependent variable is winsorized 1 percent in each tail. Standard errors clustered by state. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.9: Employment Elasticities with respect to Mortgage Supply

	Dependent variables 2007-2010:				
	Δ Constr. Emp Coef./SE	Δ Fin Emp Coef./SE	Δ Total Emp Coef./SE	Δ Other Emp Coef./SE	Δ Nontr. Emp Coef./SE
Δ Mortgage Credit 2007-2010	0.514*** (0.09)	0.390*** (0.04)	0.114*** (0.03)	0.038 (0.04)	0.032 (0.09)
All other controls	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.64	0.15	0.47	0.32	0.36
First stage F stat	27.17	25.87	27.16	26.81	25.84
Observations	1001	1026	1027	1027	1025

This table shows the effects of changes in mortgage credit, when instrumented using the nonlocal lending shock, on changes in local outcomes for the largest 850 U.S. counties. The nonlocal lending shock measures the exposure of counties to lender shocks (see text for details). All regressions include region fixed effects and all other observed characteristics of localities used in the other tables in the paper. Observations weighted by the number of employed workers in 2006. The dependent variable is winsorized 1 percent in each tail. Standard errors clustered by state. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.10: Elasticity of Construction Employment with Housing Supply Interaction

	Dependent variables 2007-2010:			
	Δ Constr. Emp Coef./SE	Δ Constr. Emp Coef./SE	Δ Permits Coef./SE	Δ Permits Coef./SE
$\Delta\widehat{Credit}$ 2007-10	0.470***	0.461***	1.011***	1.029***
	(0.09)	(0.10)	(0.24)	(0.26)
$\Delta\widehat{Credit}$ 2007-10 × Elasticity		0.081**		0.184
		(0.03)		(0.13)
All other controls	Yes	Yes	Yes	Yes
R-squared	0.70	0.70	0.56	0.56
First stage F stat	19.61	10.10	18.37	9.25
Observations	550	550	524	524

This table shows the effects of changes in mortgage credit over 2007-2010, interacted with the housing supply elasticity of Saiz (2010), on changes in construction employment and permit issuance during the recession. All regressions include region fixed effects and all other observed characteristics of localities used in the other tables in the paper (Table 1.3). The nonlocal lending shock measures the exposure of counties to lender shocks as defined in equation 1.2. Observations weighted by the number of employed workers in 2006. Standard errors clustered at the division level. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.11: OLS Estimation Results

	Dependent variables:				
	Δ Permits Coef./SE	Δ Constr. Emp Coef./SE	Δ Total Emp Coef./SE	Δ Other Emp Coef./SE	Δ Nontr. Emp Coef./SE
Δ Mortgage Credit 2007-2010	0.822***	0.391***	0.141***	0.121***	0.106***
	(0.08)	(0.06)	(0.01)	(0.01)	(0.02)
All other controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.49	0.65	0.47	0.33	0.37
Adj. R-squared	0.48	0.64	0.46	0.32	0.35
Observations	951	1001	1027	1027	1025

This table shows that the OLS coefficients when regressing changes in outcome variables (e.g. home permits) on changes in mortgage credit at the county-level over 2007-2010 while controlling for all prerecession county characteristics listed in Table 1.3. Observations weighted by the number of employed workers in 2006. Standard errors clustered at the division level. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.12: Elasticities With IV Constructed Using 2000-2002 Shares

	Dependent variables 2007-2010:				
	Δ Permits Coef./SE	Δ Constr. Emp Coef./SE	Δ Total Emp Coef./SE	Δ Other Emp Coef./SE	Δ Nontr. Emp Coef./SE
Δ Mortgage Credit 2007-2010	0.985***	0.394**	0.106**	0.037	-0.075
	(0.20)	(0.17)	(0.05)	(0.06)	(0.14)
All other controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.49	0.65	0.47	0.31	0.32
First stage F stat	14.68	15.78	15.91	16.40	16.42
Observations	951	1001	1027	1027	1025

This table shows the effects of changes in mortgage credit over 2007-2010, when instrumented using the nonlocal lending shock based on 2000-2002 market shares (as opposed to the baseline measure which uses 2005-2007 shares); see equation 1.7. All regressions include region fixed effects and all other observed characteristics of localities used in the other tables in the paper (Table 1.3). The nonlocal lending shock measures the exposure of counties to lender shocks as defined in equation 1.2. Observations weighted by the number of employed workers in 2006. Standard errors clustered at the division level. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.13: 'Placebo' Regressions

	Δ Total 90-92 Coef./SE	Δ Total 00-03 Coef./SE	Δ Constr 90-92 Coef./SE	Δ Constr 00-03 Coef./SE
Δ Mortgage Credit 2007-2010	-0.082**	0.047	0.080	0.175
	(0.04)	(0.08)	(0.17)	(0.13)
All other controls	Yes	Yes	Yes	Yes
Observations	972	973	972	971

This table reports results from 'placebo' regressions over the previous two recessions. The dependent variables are in percent change over 1990-1992 and 2000-2003. All regressions include region fixed effects and all other observed characteristics of localities used in the other tables in the paper (Table 1.3). Observations weighted by the number of employed workers in 2006. Standard errors clustered at the division level. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.14: Elasticity Estimates Including Failed Lenders

	Dependent variables 2007-2010:				
	Δ Permits Coef./SE	Δ Constr. Emp Coef./SE	Δ Total Emp Coef./SE	Δ Other Emp Coef./SE	Δ Nontr. Emp Coef./SE
Δ Mortgage Credit 2007-2010	1.168*** (0.16)	0.509*** (0.10)	0.076* (0.04)	0.001 (0.05)	0.007 (0.11)
All other controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.47	0.64	0.46	0.29	0.35
First stage F stat	31.91	27.76	28.68	28.48	27.28
Observations	951	1001	1027	1027	1025

This table shows the effects of changes in mortgage credit over 2007-2010, when instrumented using the nonlocal lending shock, including large institutions who filed for bankruptcy over 2005-2010: American Home Mortgage, New Century Financial, IndyMac, Fremont Investment, WMC Mortgage, Lehman, Ameriquest, Option One, First Magnus, and Taylor, Bean, & Whitaker Mortgage. All equations include region fixed effects and all other observed characteristics of localities used in the other tables in the paper (Table 1.3). The nonlocal lending shock measures the exposure of counties to lender shocks as defined in equation 1.2. Observations weighted by the number of employed workers in 2006. Standard errors clustered at the division level. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.15: Elasticity Estimates with Additional Controls

	Dependent variables 2007-2010:				
	Δ Permits Coef./SE	Δ Constr. Emp Coef./SE	Δ Total Emp Coef./SE	Δ Other Emp Coef./SE	Δ Nontr. Emp Coef./SE
Δ Mortgage Credit 2007-2010	1.031***	0.541***	0.126***	0.046	0.028
	(0.16)	(0.10)	(0.04)	(0.04)	(0.08)
All other controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.50	0.64	0.48	0.32	0.38
First stage F stat	22.25	24.70	24.21	23.70	23.06
Observations	951	1001	1027	1027	1025

This table shows the effects of changes in mortgage credit over 2007-2010, when instrumented using the nonlocal lending shock, on local outcomes. These regressions include squared and cubed terms for household debt-to-income, the local fraction of subprime borrowers, and the runup in home prices over 2003-2006. All equations include region fixed effects and all other observed characteristics of localities used in the other tables in the paper (Table 1.3). The nonlocal lending shock measures the exposure of counties to lender shocks as defined in equation 1.2. Observations weighted by the number of employed workers in 2006. Standard errors clustered at the division level. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 1.16: Elasticity Estimates Including Changes in Small Business Lending (CRA)

	Dependent variables 2007-2010:				
	Δ Permits Coef./SE	Δ Constr. Emp Coef./SE	Δ Total Emp Coef./SE	Δ Other Emp Coef./SE	Δ Nontr. Emp Coef./SE
Δ Mortgage Credit 2007-2010	1.037***	0.516***	0.111***	0.033	0.030
	(0.17)	(0.10)	(0.04)	(0.05)	(0.10)
All other controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.48	0.64	0.47	0.31	0.36
First stage F stat	39.75	41.11	41.62	40.64	39.47
Observations	951	1001	1027	1027	1025

This table shows the effects of changes in mortgage credit over 2007-2010, when instrumented using the nonlocal lending shock, on local outcomes. I include changes in small business lending over 2007-2010 obtained from the Community Reinvestment Act. All equations include region fixed effects and all other observed characteristics of localities used in the other tables in the paper (Table 1.3). The nonlocal lending shock measures the exposure of counties to lender shocks as defined in equation 1.2. Observations weighted by the number of employed workers in 2006. Standard errors clustered at the division level. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Chapter 2

Property Investors and the Housing Boom and Bust

2.1 Introduction

This paper argues that real estate investors – existing homeowners acquiring additional properties – played a prominent role in generating the boom-bust dynamics in economic activity observed over 2003-2010. The literature has already documented that investor activity was economically meaningful and far from anecdotal during the housing boom. Bhutta (2015) finds that the contribution of property investors to new mortgage debt in the mid 2000s exceeded in both levels and growth rates that of first-time home buyers including subprime. In fact, in the peak boom years of 2003-2006 when second and third-home buying flourished, the home ownership rate barely budged from 68.3 to 68.8 percent. Investors might have helped precipitate the bust, too. After controlling for various loan characteristics, Haughwout et al. (2011) find that property investors were more likely to default over 2007-2010.

The main contribution of this paper is to quantify the extent to which investor activity contributed to the run-up and subsequent decline in mortgage credit, home prices, and employment on a broad subset of local economies. This contributes to the lively debate on the drivers of the housing boom. It is broadly agreed that subprime and low-income borrowers levered up in the boom (Mian and Sufi 2009; Demyanyk and Hemert 2011; Gerardi, Shapiro, and Willen 2008). Growing evidence finds that prime and higher-income borrowers also contributed to the run-up in household debt (Adelino, Schoar, and Severino 2016; Foote, Loewenstein, and Willen 2016; Albanesi, Giorgi, and Nosal 2017). The findings here contribute to that evidence by emphasizing the relatively unexplored role played by property investors.[1]

Increases in investor activity in the boom years could have led to increases in housing demand, and thereby increases in home prices and construction. In the bust, the effects could have turned contractionary. Property investors were highly levered and experienced higher default rates in the recession than first time homeowners (Haughwout et al. 2011; Bayer et al. 2011). Moreover, excessive home building would have led to inefficient land use in the recession difficult to overcome due to irreversibility constraints (Boldrin et al. 2016).

The main identification challenge is of reverse causality – surging investor activity could have driven home price appreciation, but it is also plausible that expected home price appreciation drove investor activity. I measure investor activity at the county level as the share of mortgage originations for non owner-occupied housing from the Home Mortgage Disclosure Act (HMDA) dataset. My baseline measure is taken over 1998-2000, so it significantly predates the

[1]Financial developments have also been linked to the run-up in household debt and home prices in the U.S. such as the rise in alternative mortgages (Barlevy and Fisher 2012; Foote et al. 2008), securitization (Keys et al. 2010; Nadauld and Sherlund 2009; García 2017), and demand for mortgage derivatives from Europe (Shin 2012; Justiniano, Primiceri, and Tambalotti 2013).

peak years of the housing boom. It is therefore unlikely that the cross-sectional variation in the investor shares is driven by variation in expectations about home price appreciation. In fact, the results in the paper are robust to measuring investor activity in earlier periods such as the mid 90s.

While the investor shares are measured during the pre-boom years 1998-2000, they are an excellent predictor of cross-sectional variation in investor shares in the peak boom years. The reason is that counties with high investor activity have fixed appealing physical qualities, such as warm winters and a waterfront, as measured from the Department of Agriculture's Natural Amenities Scale. Top counties include several locations in Florida as well as the home counties of Myrtle Beach, SC, Mohave, AZ. and Maui, HI. Because of these fixed qualities, the cross-sectional variation in investor activity at the county level is highly persistent. The cross-sectional correlation is 0.88 of investor activity over 1998-2000 with activity measured over 2004-2006.[2] Investor activity is also highly correlated (0.83) with the share of vacation homes in a county, obtained from the 2000 Decennial Census. The vacation share is based on the stock of housing, and is therefore not likely influenced by short to medium-run changes in home price appreciation expectations.

Investor activity in HMDA includes buyers of vacation homes as well as 'flippers', since HMDA only distinguishes between households intending to use a property as "owner-occupied as a principal dwelling" or not. The coverage of both types is likely high in HMDA, because the majority of both second-home

[2] The correlation coefficient is 0.95 between investor activity in 1994-1996 and 1998-2010.

buyers as well as professional investors with multiple properties used mortgages in the boom.[3] Reflecting the high coverage of both types, the investor shares (over 1998-2000) are highly correlated with both the 2000 Census vacation shares, as well as measurements of speculative activity in the peak boom years. Using proprietary data on home flips – a flip is defined as the second sale of a residence within a one-year period – I find that property flip rates were higher in the peak boom years (and increased by more) in counties with high investor shares.[4]

To estimate the effects of investor activity on home prices and employment, I model the cross-sectional variation in county-level investor shares over 1998-2000 as fixed, but with potential-time varying effects from the interaction of the investor shares with year dummies. That allows for investor shares to be associated positively with higher home prices in the mid 2000s, and with lower home prices later in the decade, for example. The full model includes county and year fixed effects, as well as the interaction of a detailed set of county characteristics with year dummies. The results are robust, for instance, to controlling for the interaction of the **?** housing supply elasticity with year dummies, and so shed new light on the drivers of the housing boom and bust.

In counties with above-average investor shares, mortgage originations, home

[3] Mills, Molloy, and Zarutskie (2017) find that 68 percent of 'investor' households used mortgages to finance home purchases over 2004-2006 based on a CoreLogic dataset of county property tax assessors' records. Investors are defined as households owning three or more properties with an adjustment to distinguish from wealthy individuals that own multiple homes for personal use. This was provided upon request by Mills, Molloy, and Zarutskie (2017), who in the paper report analogous data for 2012-2014.

[4] The correlations between investor shares and flip rates and the change in flip rates are 0.52 and 0.43, respectively.

prices, and employment grew faster over 2003-2006. Counties with a 10 percentage point higher investor share (over 1998-2000), for instance, experienced on average close to 9.5 percentage point higher construction employment in 2006, with 6.5 of those percentage points explained by faster growth over 2003-2006. Higher investor shares are also associated with higher and growing levels of home prices and mortgage credit during that period.

The positive association begins to reverse in 2007. Over the next few years, counties with high investor shares over 1998-2000 crashed harder. By 2010, high investor shares were now negatively associated with lower credit issuance, home prices, and employment. In 2010, counties with a 10 percentage point higher investor share (over 1998-2000) had on average 7 percent lower construction employment. The dramatic reversal in the Great Recession years shows that 'investor counties' experienced a more pronounced boom, and a more pronounced bust.

I then attempt to quantify an answer to the question, how different would home prices and employment dynamics have been over 2003-2010, in the absence of property investment? I do so by comparing the evolution of home prices and employment against a counterfactual in which the rise in cross-sectional variation in investor shares does not help explain variation in the time series of the dependent variables (such as construction employment). Specifically, I compute fitted values from a full regression model for each dependent variable, and compare those fitted values against the counterfactual in which the coefficients are set to zero (while holding everything else constant) for the interactions of investor shares with year dummies.

In the counterfactual, mortgage credit, home prices, and construction employment grow less in the boom, and land more softly in the bust. The boom and bust are still there, reflecting the fact that other factors were important drivers, but the paths would have been smoother, particularly for construction and financial employment, where real estate intermediaries make up about a third of employees. Almost 30 percent of the rise over 2003-2006 and fall over 2007-2010 in construction and financial employment can be explained by property investment. For other employment categories (total private employment excluding construction and finance), the effects are less symmetric. Property investment is associated with a small rise in other employment in the boom years, though it can explain about a third of the employment losses in the bust. This is in line with the investment overhang hypothesis, where excessive home building in the boom creates asymmetric gains and losses in other sectors. In Boldrin et al. (2016) irreversibility constraints imply housing structures cannot be put to use by more productive industries. The zero lower bound could have also hindered the reallocation of resources to nonresidential sectors (Rognlie, Shleifer, and Simsek (Forthcoming)).

The identifying assumption is that the 1998-2000 investor shares are uncorrelated with unobserved characteristics of counties affecting boom-bust dynamics. Because high investor counties have high amenity values, these counties might be systematically different in ways that are correlated with boom-bust dynamics. To address that concern, I control for the interaction of a detailed set of county characteristics and year dummies. Moreover, I perform a useful bounding exercise by 'over-controlling' using the 2006 median ratio of household debt to income from Mian and Sufi (2014). This proxies for any unobserved local shocks that led to higher local household debt levels. I interpret these estimates as lower-bound, since controlling for 2006 debt to income also controls for local property investments made by locals (but not out-of-towners). When doing so, the qualitative results all hold, with the quantitative effects of property investment being about 30 percent smaller.

The evidence in this paper is in line with theories of the housing boom emphasizing the role played by property investors. In the housing search model of Piazzesi and Schneider (2009), it only takes a few households turning bullish to generate a housing boom. Optimists can influence prices because the volume of transactions with respect to the housing stock is relatively low.[5] The boom-bust can be protracted and deeper if bullish households find new converts, as in the search and social dynamics model of Burnside, Eichenbaum, and Rebelo (2016).[6] Short-term investment was indeed sizable, with sales of homes held for less than 3 years accounting for 42 percent of the growth in sales volume from 2000-2005 (DeFusco, Nathanson, and Zwick 2017).[7]

The main contribution of this paper is to quantify the effects of the boom and bust in property investment on economic activity over 2003-2010. This builds on previous work documenting the substantial role played by property investors in explaining the run-up in household debt, such as Haughwout et al. (2011) and Bhutta (2015) which are based on the location of the investor (Equifax)

[5] Only 6 percent of owner-occupied homes traded every year according to the American Housing Survey. In contrast, on the New York Stock Exchange, the ratio of annual volume to market capitalization is 120 percent (Piazzesi and Schneider (2009)).

[6] More than 10 TV shows were dedicated to home flipping in the mid 2000s, including Flip This House, Flip That House, and My House is Worth What?

[7] In the HMDA data, the number of loans for non owner-occupied properties increased by 63 percent from 2003-2006, compared with only 15 percent increase for owner-occupied mortgages.

rather than the investment. The HMDA data, in contrast, is based on the location of the property. Other related work includes Chinco and Mayer (2016); Gao, Sockin, and Xiong (2017); Nieuwerburgh and Favilukis (2017); Bayer et al. (2011). Using a high-frequency identification approach in a panel VAR, Chinco and Mayer (2016) find that positive shocks to investor activity led to higher home prices in the boom. They also find that out-of-town investors were less informed than locals. Their monthly transactions-level data is very detailed, but is only available for 21 MSAs. The greater geographic coverage of the HMDA data allows me to document new cross-sectional facts about investor activity such as the the close relation between HMDA investor shares, vacation shares, and flip rates. Gao, Sockin, and Xiong (2017) use state-level variation in capital gains taxes to instrument for investor activity, and also find that investor activity in the boom years is associated with declines in home prices and employment in the Great Recession. Nieuwerburgh and Favilukis (2017) solve a spatial equilibrium model of a city and find that an influx of out-of-town real estate buyers can push up construction employment and home prices, benefiting local home owners and hurting renters.[8]

2.2 Motivation

Investor counties – locations which tend to be the recipients of second home buying – experienced boom-bust dynamics in economic activity over 2003-2010.

[8]Other related work includes the growing quantitative literature assessing the extent to which different shocks can account for the stylized facts in the housing boom, such as the increase in home prices and the home-ownership rate from 2000-2005 (Justiniano, Primiceri, and Tambalotti 2015; Garriga, Manuelli, and Peralta-Alva 2012; Boldrin et al. 2016). **?** find that improved expectations of home price appreciation are necessary to explain the run-up in home prices; increases in credit supply alone do not, if they only encourage renters to become home-owners rather than buy more housing.

I measure the investor share in a county as the fraction of non owner-occupied mortgages issued in the pre-boom years of 1998-2000 from HMDA. I then divide counties into three groups: the top quartile by investor activity, the middle quartiles, and the bottom quartile. Figure 2.1 plots the average level of home prices and construction employment for each of these groups. The data are in log levels and have been indexed such that the log levels equal one in the year 2000. The figure shows that the 3 groups exhibit common trends until about 2003; from 2003-2006 investor counties experience a boom; and from 2007-2010 investor counties experience a sharp contraction. By 2010, the groups roughly seem to be trending similarly again.

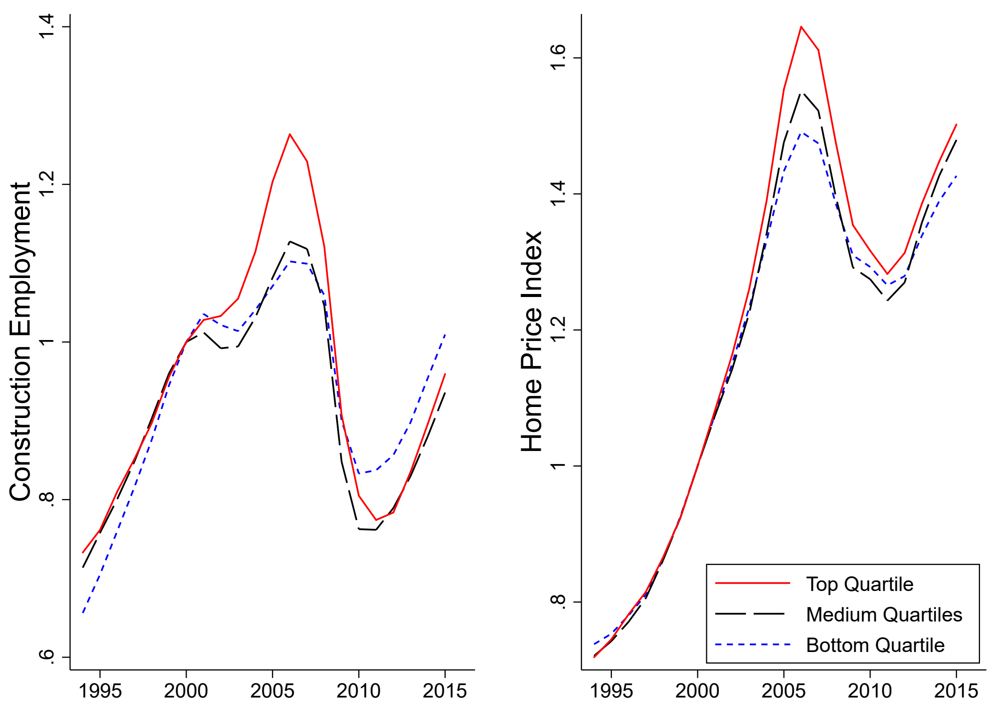
As further motivation, I show that counties with high investor shares (1998-2000) were more likely to experience a housing boom over 2003-2006. I compute the historical mean and standard deviation for annual growth in home prices for each county using the FHFA Home Price Index, which is available back to the 1980s for most of all the largest 500 counties in the country. I define the county-level indicator $Boom_i$ as equal to 1 if yearly growth rates between 2003-2006 exceeded twice the standard deviation of growth rates plus the historical mean. Of the largest 500 counties, close to 20 percent of counties fit the bill as having experienced a housing boom.

Columns 1 and 2 of Table 2.1 provide coefficient estimates of the following probit model:

$$P(Boom_i = 1) = \Phi(\alpha_0 + \alpha_1 InvestorShare_{i,98-00} + \alpha_2 X_{i,2000} + \upsilon_i)$$

where $Boom_i$ is an indicator variable for whether county i experienced a home

Figure 2.1: Construction Employment and Home Price Boom-Bust



All series are in log levels with the value in year 2000 set to 1. Counties are divided into quartiles by the share of real estate investor activity, measured over 1998-2000. Source: Quarterly Census of Employment and Wages, and Zillow Home Price Index

price boom; the main explanatory variable of interest is the investor share measured over 1998-2000; X_i are other county characteristics in 2000 obtained from the Census: median income and home values in dollars, the fraction of college-educated, senior citizens, white, poor, housing units with an outstanding mortgage, and home values exceeding the conforming loan limit; and υ_i is an error term. The investor share is standardized and Table 2.1 reports marginal effects when holding all explanatory variables constant at their means. Therefore the Table coefficients have the interpretation of the increase in the likelihood of a county experiencing a housing boom over 2003-2006 from having an investor share one standard deviation above the mean. Errors are clustered at the state-level to allow for arbitrary correlation of shocks within states.

Table 2.1: The Effect of Investor Activity on the Likelihood of Experiencing a Housing Boom

	Probit		ΔHome Price 2003-2006	
Investor Share 1998-2000	.074*	.085***	.281**	.353***
	(.044)	(.035)	(.112)	(.098)
Other Controls	No	Yes	No	Yes
$E[Boom]$	.18	.18		
$E[Boom : InvShare_{P90} - InvShare_{P10}]$	.13	.14		
Observations	497	497	497	497

The dependent variable in Columns 1-2 is an indicator for whether the county experienced a boom in home prices over 2003-2006, as defined as in text. The Investor Share is the fraction of non owner-occupied mortgages for home purchase over 1998-2000. The dependent variable in Columns 3-4 is percent change in home prices over 2003-2006. Columns 2 and 4 contain a full set of county characteristics in 2000 obtained from the Census: median income and home values in dollars, the fraction of college-educated, senior citizens, white, poor, housing units with an outstanding mortgage, and home values exceeding the conforming loan limit. Columns 1 and 2 report marginal coefficients from probit evaluated at the mean of all explanatory variables. Columns 3 and 4 report OLS coefficients. All standard errors clustered by state.

The coefficient on the investor share indicates that counties with higher investor activity were more likely to experience a housing boom over 2003-2006. Column 1 provides results from the bivariate specification while Column 2 includes the full set of pre-boom county controls. With the full set of controls, a 1 standard deviation increase in the investor share is associated with an 8.5 percentage point increase in the likelihood of a county experiencing a housing boom. For a county in the 10th percentile of the distribution of investor share, the likelihood of having a housing boom is 9 percent, holding all other variables constant at their means. When changing the investor share to the 90th percentile, the likelihood increases by 14 percentage points to 23 percent – a greater than doubling in the likelihood of experiencing a boom episode. In Columns 3 and 4, the dependent variable is the percent change in home prices over 2003-2006, and the models are estimated via OLS. The coefficient on the investor share is positive and significant providing complementary evidence that investor counties experienced faster growth in home prices.

The rest of the paper discusses the data used in the paper (Section 2.3) and documents investor counties experienced higher speculative activity (flip rates) in the boom years. Section 2.4 shows the increase in speculation was accompanied by a boom and bust in labor and housing markets. Moreover, the documented boom-bust dynamics are associated specifically with cross-sectional variation in investor activity, and not other dimensions previously explored in the literature such as the housing supply elasticity **?** or household leverage (Mian and Sufi (2014)).

2.3 The Geography of Property Investment

The cross-sectional variation in property investment is to a large extent driven by variation in the physical appeal of locations. Because these qualities are mostly fixed, the investor shares are highly correlated year-on-year. There is significant cross-sectional variation in investor activity (measured over 1998-2000), with the bottom 10 percent of counties having less than 4.3 percent investor shares, and the top 10 percent with over 16.3 percent shares of investor activity. Counties with high investor shares are located in areas with appealing features such as a waterfront. Examples include various counties in FL, and the home counties of Myrtle Beach, SC, Mohave, AZ. and Maui, HI.

I measure investor shares over 1998-2000 to avoid the potential reverse causality concern that cross-sectional variation in expected home appreciation in the boom years drove the geographic variation in property investment.[9] As evidence that the physical qualities of locations mostly explain cross-sectional variation in investor activity, Figure 2.2 shows that investor shares (over 1998-2000) are highly correlated with the share of vacation homes in a county. The correlation coefficient is 0.83, and the vacation shares are obtained from the 2000 Decennial Census. The vacation share measure is based on the stock of housing and so are not likely to be influenced by any recent trends.

Investor shares over 1998-2000 are an excellent predictor of investor shares in the peak boom years of 2004-2006 (and other periods as well). Figure 2.3 shows the correlation between investor shares measured over 1998-2000 and 2004-2006 is 0.88. The correlation is also close to one for other periods. Counties with high investor shares tend to have appealing physical qualities, such as warm and sunny winters, as well as proximity to water. The Natural Amenities dataset of the Department of Agriculture compiles six measures of the physical qualities of locations – temperatures in January and July, hours of sunlight in January, humidity in July, a topographic measure ranging from plains to mountains, and the fraction of water area in a county. These characteristics explain between 20 and 40 percent of the cross-sectional variation in investor shares over 1998-2000

[9]Measuring the investor share at an earlier period such as 1994-1996 leads to nearly identical results, with the correlation coefficient equal to 0.95.

and 2004-2006, as shown in Table 2.2. This is particularly true for the very top investor counties, which essentially all have a sizable waterfront.

Figure 2.2: Vacation Counties

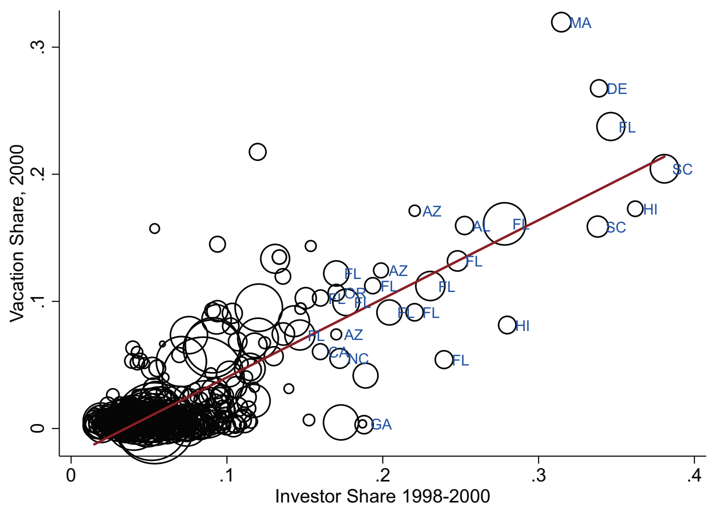


Figure plots the share of vacation homes in the 2000 Decennial Census versus investor shares in 2000-2003. Observations weighted by the number of home sales in 2005-2006.

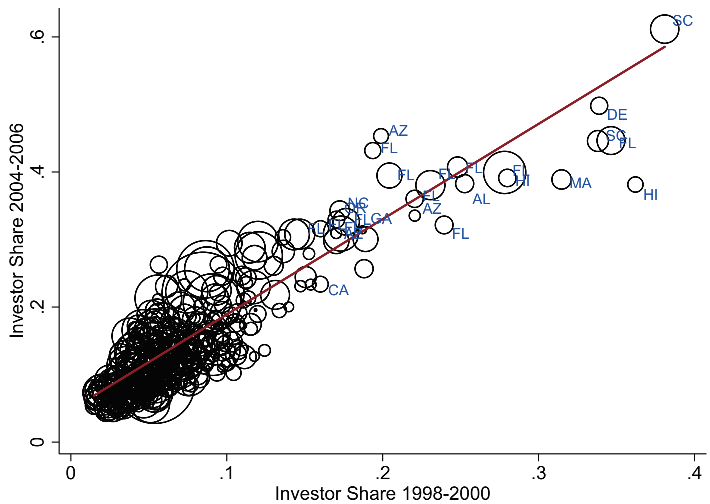
Investor activity in HMDA would include different types of investors, such as buyers of vacation homes, rental properties, and flips, since HMDA only distinguishes between households intending to use a property as "owner-occupied as a principal dwelling" or not. Coverage depends on the extent to which different investor types use mortgages. Coverage is likely reasonably high for all of the investor types, with some variation. Mills, Molloy, and Zarutskie (2017) document the extent to which different buyer types tend to use mortgages when buying homes, using CoreLogic data on property transactions and county property tax assessors' records. About 85 percent of non-investor households used mortgages to finance home purchases in 2004-2006. This group would be most associated with buyers of second homes (such as vacation homes), and so are

Table 2.2: Investor Shares and Natural Amenities

	Investor Share $2004-2006$ Coef./SE	*Investor Share* $1998-2000$ Coef./SE
Temperature in January	0.38**	0.30***
	(0.15)	(0.10)
Water Area	0.18**	0.18**
	(0.07)	(0.07)
Temperature in July	0.16	-0.01
	(0.14)	(0.12)
Hours of Sunlight in January	0.09	0.11
	(0.11)	(0.09)
Humidity in July	-0.18**	-0.07
	(0.08)	(0.07)
Topography	-0.23***	-0.16***
	(0.08)	(0.05)
R-squared	0.39	0.23
# Counties	418	418

The dependent variables are the county-level investor shares, over 2004-2006 in Column 1, and 1998-2000 in Column 2. The explanatory variables are physical characteristics of localities obtained from the Department of Agriculture's Natural Amenities dataset. All variables are standardized. Observations are weighted by the number of home sales in 2005-2006.

Figure 2.3: Stable Classification of Investor Counties



Source: HMDA. Figure plots investor shares in 2005-2006 against investor shares in 2000-2003. Observations weighted by the number of home sales in 2005-2006.

likely the best represented in the HMDA data. The analogous statistic is also high for 'investors' – about 68 percent of home-purchases for this category where mortgage-financed. Investors in Mills, Molloy, and Zarutskie (2017) are defined as households owning three or more properties with an adjustment to distinguish from wealthy individuals that own multiple homes for personal use.[10]

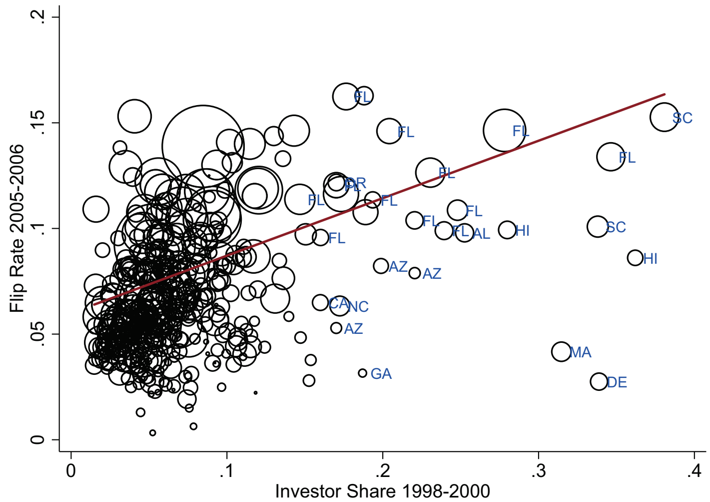
Consistent with HMDA coverage of potential flippers being reasonably high, the investor share (over 1998-2000) is positively associated with measures of speculative activity in the peak boom years. I obtain flip rate data from RealyTrac – a flip is defined as a second sale of a residence within a one-year period. The flip rate is the total number of flips in a county in a given year divided by the total number of home sales, and the data are based on public deed records.

[10]Mills, Molloy, and Zarutskie (2017) document that in the recovery period (2012-2014) property investments are now less household and mortgage-driven, with corporate investors accounting for a larger share of property investment.

Three years of data were acquired for 2001, 2005, and 2006.

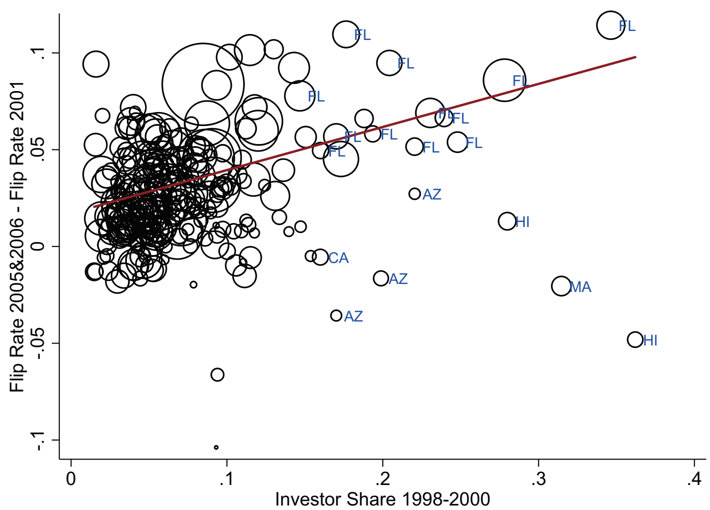
Counties with a 10 percent higher share of investor activity experienced on average 2.7 percentage point higher flip rates over 2005-2006 (Figure 2.4). Flip rates also increased more in investor counties. A 10 percent higher share of investor activity is associated with a 2.2 percentage point increase in flip rates between 2001 and 2005-2006 (Figure 2.5). The correlations of investor activity over 1998-2000 and flip rates over 2005-2006 and the change in flip rates are 0.53 and 0.42, respectively.

Figure 2.4: Flip Rates vs Investor Activity



Source: RealtyTrac. Figure plots flip rates in 2005-2006 against investor shares in 1998-2000. Observations weighted by the number of home sales in 2005-2006. Data available for 418 of the 500 largest counties.

Figure 2.5: Change in Flip Rates vs Investor Activity



Source: RealtyTrac. Figure plots the increase in flip rates from 2001 to 2005-2006 against investor shares in 1998-2000. Observations weighted by the number of home sales in 2005-2006. Data available for 269 of the 500 largest counties.

There was under-reporting of investment activity in the HMDA dataset in the time series, though this is not likely to be a major concern for the pre-boom cross-sectional measurement of investor activity. All else equal, primary residence mortgages tend to have more favorable loan terms, while sales of primary residences are taxed at lower rates. Partly because of that, owner-occupancy was under-reported particularly in the peak boom years (Elul and Tilson 2015, Piskorski, Seru, and Witkin 2015, Mian and Sufi 2015). To the extent that the loosening of documentation standards enabled misreporting in the peak boom years, measurement error is likely to be less of an issue before the 2000s. The main measurement of investor activity used in this paper is over 1998-2000, and is cross-sectional, so measurement concerns are likely to be minor. As evidence for that, the investor activity measure is highly correlated with related county-level indicators obtained from independent datasets. For instance, the correlation coefficient is 0.83 between investor shares and the share of vacation homes from the Census.

2.4 Investor Activity and Boom-Bust Dynamics

Property investment surged in the peak boom years 2003-2006. I hypothesize that this surge in property investment had a stronger effect on counties which traditionally were the recipients of investor activity. I model the cross-sectional variation in investor shares as fixed and with potential time-varying effects through the interaction of county-specific investor shares with year dummies. To the extent that the investor shares measured over 1998-2000 are not systematically associated with other, unobserved local factors explaining boom-bust dynamics, the interactions will reveal the effects of investor activity on local economic activity. Specifically, I estimate the following fixed effects model:

$$Y_{it}^{j} = \alpha_i + \tau_t + \beta_t(Investor\ Share_i \times \tau_t) + \phi_t(Z_i \times \tau_t) + \epsilon_{it} \tag{2.1}$$

for counties i and years t. The dependent variables Y_{it}^{j} are in log-levels and include (indexed by j) the flow of new mortgage credit (originations), home prices, construction employment, financial employment, and 'other' employment defined as total private minus construction and financial employment. Each model is estimated separately. The specification includes a full set of county α_i and year fixed effects τ_t. Data for mortgage originations, employment categories, and home prices come from the Home Mortgage Disclosure Act, the Quarterly Census of Employment and Wages, and Zillow Research. All series run from 1994 to 2015 with the exception of the Zillow home price index which starts in 1996.

The parameters of interest are β_t – differences in the level of the dependent variable explained by variation in the investor share on a year-specific basis. The effects are allowed to vary by year over the 2003-2012 period, in order to investigate trends. For example, having a high investor share could be associated with higher home prices in the mid 2000s, but lower home prices in late 2000s. Standard errors are clustered at the county-level to account for serial correlation in the residuals. The dependent variables are in levels, to allow for the explanatory variables having potentially persistent effects on the level of economic activity. The qualitative conclusions are the same when estimating in first differences or with lags, as shown in the Robustness section.

One concern is that variation in the investor shares is correlated with other characteristics of localities, which may themselves be associated with boom-bust dynamics. For example, investor counties tend to be lower income. To ensure that β capture only the effects of variation in investor activity, I control for the interaction of year dummies with other cross-sectional local characteristics Z_i all measured in a pre-boom period, such as median home values, household income, the fraction of the population that is college-educated, poor, have an open mortgage, identify as white, are 55 years old or older, live in an urban area (all obtained from the 2000 Decennial Census), and the manufacturing share of employment in 2000 from the QCEW. I also control for pre-boom trends in the performance of local economies, specifically, the growth in home prices, construction employment, and other employment over 1998 and 2003.[11] These pre-boom characteristics are also interacted with the year dummies.

Table 2.3 shows results for the baseline specification. Mortgage originations, home prices, and employment grew faster in investor counties over 2003-2006. For example, counties with a 10 percentage point higher investor share experienced on average close to 9.5 percentage point higher construction employment in 2006, with 6.3 of those percentage points explained by faster growth over 2003-2006. Across models, higher investor shares are associated with higher and growing levels of economic activity during that period.

The positive association begins to reverse in 2007. Over the next few years,

[11] The results are also robust to controlling for the housing supply elasticity of Saiz (2010), though that measure is available only for a smaller sample of counties. The R-squared in a population-weighted (unweighted) regression of the elasticity against the investor share is 0.002 (0.004).

investor counties crashed more deeply. By 2010, high investor shares were then negatively associated with lower credit issuance, home prices, and employment. In 2010, counties with a 10 percentage point higher investor share had on average 6.8 percent lower construction employment. The dramatic reversal in the Great Recession years shows that investor counties experienced a more pronounced boom, and a more pronounced bust.

The identifying assumption is that the 1998-2000 investor shares are uncorrelated with unobserved characteristics of counties affecting boom-bust dynamics. For example, unobserved productivity shocks could be correlated with the investor shares. In the absence of a direct measure of local productivity shocks, I control for cross-sectional variation in median household debt to income in 2006. Local positive productivity shocks perceived to be permanent would induce households in the county to lever up. Mian and Sufi (2011) have documented that median household debt to income in 2006 is an excellent predictor of boom-bust dynamics.

The measure of household debt to income, based on the New York Fed's Consumer Credit Panel/Equifax, is comprehensive and local – it includes mortgages as well as other household debt such as auto loans, and is based on the primary residence of the households. The run-up in household debt incurred by local investors buying local properties would be captured in the debt-to-income measure. Therefore, including debt to income as a regressor would amount to over-controlling in the sense that it would also partly capture local second-home investment. This is helpful for identification purposes, but could lead to underestimates, since the estimates would reflect only property investment driven by

Table 2.3: Investor Activty and Boom-Bust Dynamics

	Dependent variables:				
	$Originations_{it}$	$Home\ Prices_{it}$	$Constr.\ Emp_{it}$	$Fin.\ Emp_{it}$	$Other\ Emp_{it}$
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
β_{2003}	0.73***	0.21**	0.32***	0.05	0.06**
	(0.14)	(0.10)	(0.06)	(0.07)	(0.03)
β_{2004}	1.06***	0.22**	0.51***	0.11*	0.07**
	(0.16)	(0.11)	(0.08)	(0.07)	(0.03)
β_{2005}	1.30***	0.59***	0.85***	0.20***	0.10***
	(0.20)	(0.16)	(0.11)	(0.08)	(0.03)
β_{2006}	0.66***	0.67***	0.94***	0.24***	0.13***
	(0.21)	(0.20)	(0.13)	(0.09)	(0.04)
β_{2007}	-0.09	0.37**	0.73***	0.25***	0.09**
	(0.19)	(0.16)	(0.12)	(0.09)	(0.04)
β_{2008}	-0.76***	0.02	0.12	0.14	-0.01
	(0.20)	(0.13)	(0.11)	(0.10)	(0.04)
β_{2009}	-0.88***	-0.39***	-0.45***	-0.04	-0.13***
	(0.20)	(0.14)	(0.13)	(0.10)	(0.05)
β_{2010}	-0.86***	-0.72***	-0.69***	-0.04	-0.16***
	(0.20)	(0.18)	(0.14)	(0.11)	(0.06)
β_{2011}	-0.67***	-0.92***	-0.88***	-0.09	-0.15**
	(0.20)	(0.19)	(0.15)	(0.12)	(0.06)
β_{2012}	-0.60***	-0.85***	-0.88***	-0.12	-0.14**
	(0.19)	(0.18)	(0.15)	(0.13)	(0.06)
County, Year FE	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes
Within R-squared	0.78	0.83	0.59	0.35	0.56
# Observations	17283	10320	17864	18106	18084
# Counties	823	516	812	823	822

The table reports estimates based on the model

$$Y_{it}^{j} = \alpha_i + \tau_t + \beta_t(Investor\ Share_i \times \tau_t) + \phi_t(Z_i \times \tau_t) + \epsilon_{it}$$

where β_t is the coefficient associated with the interaction of the investor share over 1998-2000 and a year t dummy variable Regressions include county fixed effects and time fixed effects for the sample of largest counties for which data is available. Additional controls include the interaction of year dummies with county characteristics Z_i: 2000 median home values, household income, the fraction of the population that is college-educated, poor, have an open mortgage, identify as white, are 55 years old or older, live in an urban area, and the manufacturing share of employment in 2000 from the QCEW. Standard errors clustered at the county-level.

out-of-town investors. Therefore, I interpret the coefficient estimates, when controlling for cross-sectional variation in household debt to income, as providing a lower bound.

Table 2.4 shows results for the specification that controls for household debt to income. The qualitative patterns are the same. Investor counties grew faster in the boom years, and crashed harder in the years of the Great Recession. The magnitudes of the boom and bust explained by investor activity are smaller than in the baseline case, as would be expected, since local property investments in the boom years made by locals would be captured in the debt-to-income measure. Because the qualitative results still hold, and the magnitudes are comparable, this suggests that out-of-town investors accounted for a large share of the investments.

2.4.1 Aggregate Implications

How different would home price and employment dynamics been over 2003-2012 in the absence of property speculation? This section attempts to provide an answer. Specifically, I compare the evolution of home prices and employment against a counterfactual in which cross-sectional variation in investor activity is not helpful in explaining time-series variation.

Denote $\widehat{Y_{it}^{j}}$ as the fitted values of equation 2.1 – the fitted values from the full model of dependent variable j for county i and year t:

$$\widehat{Y_{it}^{j}} = \hat{\alpha}_i + \hat{\tau}_t + \hat{\beta}_t(Investor\ Share_i \times \tau_t) + \hat{\phi}_t(Z_i \times \tau_t) \tag{2.2}$$

These fitted values contrast with the counterfactual $\widehat{Y_{it}^{j,CF}}$ where, all else equal, cross-sectional variation in investor activity does not help explain dynamics in

Table 2.4: Investor Activity and Boom-Bust Dynamics – Controlling for Cross-Sectional Variation in 2006 Household Debt to Income

	Dependent variables:				
	$Originations_{it}$	$Home\ Prices_{it}$	$Constr.\ Emp_{it}$	$Fin.\ Emp_{it}$	$Other\ Emp_{it}$
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
β_{2003}	0.44***	0.05	0.18***	-0.08	0.02
	(0.13)	(0.09)	(0.06)	(0.08)	(0.03)
β_{2004}	0.72***	0.02	0.24***	0.02	0.03
	(0.15)	(0.09)	(0.08)	(0.08)	(0.03)
β_{2005}	0.89***	0.24*	0.48***	0.12	0.05*
	(0.21)	(0.14)	(0.10)	(0.09)	(0.03)
β_{2006}	0.60***	0.25	0.51***	0.16*	0.07*
	(0.23)	(0.19)	(0.14)	(0.10)	(0.04)
β_{2007}	0.36*	0.05	0.44***	0.21**	0.04
	(0.19)	(0.17)	(0.11)	(0.10)	(0.04)
β_{2008}	-0.33	-0.07	0.09	0.20*	-0.01
	(0.21)	(0.14)	(0.12)	(0.11)	(0.05)
β_{2009}	-0.77***	-0.28*	-0.19	0.07	-0.07
	(0.21)	(0.15)	(0.14)	(0.11)	(0.05)
β_{2010}	-0.67***	-0.55***	-0.34**	0.15	-0.06
	(0.20)	(0.17)	(0.14)	(0.12)	(0.05)
β_{2011}	-0.55***	-0.53***	-0.53***	0.11	-0.04
	(0.21)	(0.17)	(0.15)	(0.13)	(0.06)
β_{2012}	-0.47**	-0.47***	-0.53***	0.06	-0.03
	(0.20)	(0.17)	(0.15)	(0.14)	(0.06)
County, Year FE	Yes	Yes	Yes	Yes	Yes
2006 Debt to Income	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes
Within R-squared	0.78	0.84	0.60	0.35	0.56
# Observations	17283	10320	17864	18106	18084
# Counties	823	516	812	823	822

The table reports estimates based on the model

$$Y_{it}^{j} = \alpha_i + \tau_t + \beta_t(Investor\ Share_i \times \tau_t) + \phi_t(Z_i \times \tau_t) + \epsilon_{it}$$

where β_t is the coefficient associated with the interaction of the investor share over 1998-2000 and a year t dummy variable. Regressions include county fixed effects and time fixed effects for the sample of largest counties for which data is available. Additional controls include the interaction of year dummies with county characteristics Z_i: 2006 median household debt to income, 2000 median home values, household income, the fraction of the population that is college-educated, poor, have an open mortgage, identify as white, are 55 years old or older, live in an urban area, and the manufacturing share of employment in 2000 from the QCEW. Standard errors clustered at the county-level.

the dependent variable – specifically, where $\widehat{\beta}_t = 0$ for all years between 2003-2012.

$$\widehat{Y_{it}^{j,CF}} = \widehat{\alpha}_i + \widehat{\tau}_t + \widehat{\phi}_t(Z_i \times \tau_t) = \widehat{Y_{it}^{j}} - \widehat{\beta}_t(Investor\ Share_i \times \tau_t) \tag{2.3}$$

Figure 2.6 provides a visual comparison of the evolution of aggregate home prices, construction, financial, and other employment for the fitted values of the full model (equation 2.2), against the counterfactual in which $\widehat{\beta}_t = 0$ from equation 2.3. Each county-specific series is rescaled to equal 1 in the year 2000. The aggregate series are then obtained by taking the population-weighted average across counties. The aggregate version of the data and the fitted values from the model are almost indistinguishable, so the figure only includes the fitted values in addition to the counterfactual.

In the counterfactual economic activity grew more slowly in the boom, and declines less precipitously in the bust. This shows that home prices and employment (especially in construction and finance) would have been less volatile in the absence of the rise and collapse in speculative investment. That said, the dynamics in the fitted values of the full model and the counterfactual are fairly similar, with economic activity growing over 2003-2006, and contracting in the following years, reflecting the fact that other factors other than the property investment explain the majority of boom-bust dynamics.

I define the percent variation explained by investor activity as the difference between the changes in the fitted values of the full model and the changes in the counterfactual, relative to the observed changes in the data. In particular, let $\Delta Y^j_{boom}, \widehat{\Delta Y^j_{boom}}, \widehat{\Delta Y^{j,CF}_{boom}}$ stand for the change in the observed data, fitted values, and counterfactual, respectively, for dependent variable j over 2003-2006. The

Figure 2.6: Economic Activity vs Counterfactual

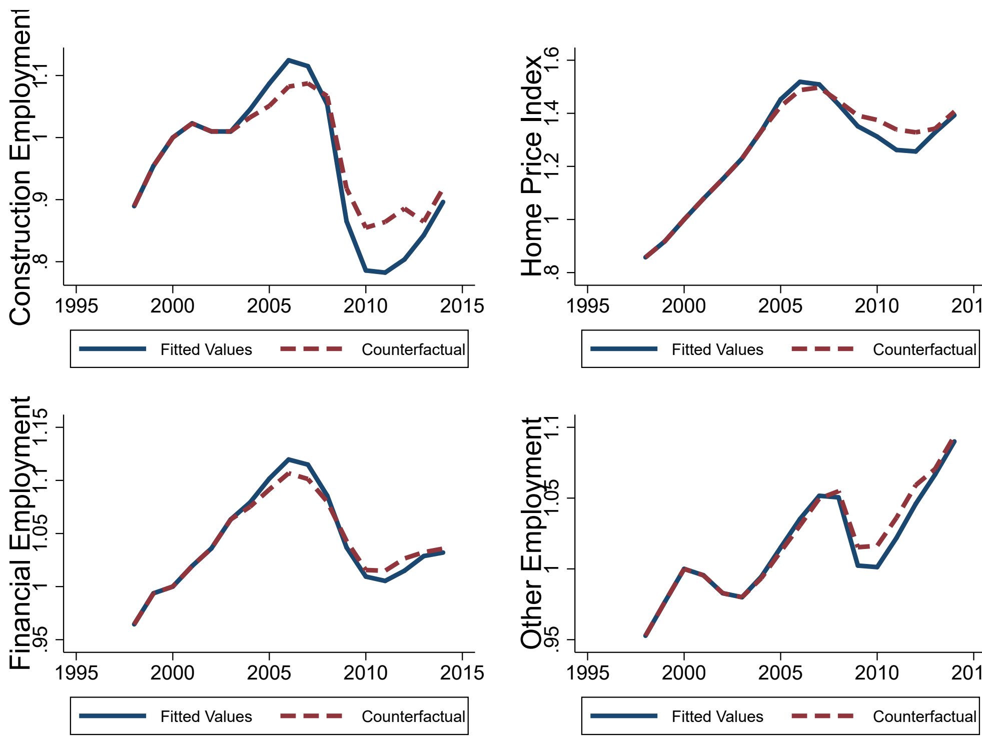


Figure plots fitted values based on equation 2.2 against a counterfactual in which variation in investor shares does not help explain growth in the boom or the collapse in the bust for each dependent variable (equation 2.3).

percent variation explained by investor activity for dependent variable j over 2003-2006 is given by

$$\frac{\widehat{\Delta Y^{j}_{boom}} - \widehat{\Delta Y^{j,CF}_{boom}}}{\Delta Y^{j}_{boom}} \tag{2.4}$$

with an analogous definition for the bust period of 2007-2010.

Table 2.5 quantifies the extent to which investor activity explains variation in home prices, construction, financial, and other employment in the boom and bust periods. Two sets of estimates are provided: first, for the baseline

Table 2.5: Percent of Observed Changes in Economic Activity Explained by Investor Activity

	Construction Emp	Home Price Index	Financial Emp	Other Emp
Baseline estimates				
2003-2006:	36.95	10.40	20.28	8.14
2007-2010:	29.51	36.11	15.79	35.70
Lower-bound estimates: Controlling for variation in 2006 household debt to income				
2003-2006:	20.02	4.75	29.95	6.85
2007-2010:	16.00	18.57	2.7	14.37

This table computes the percent of the observed changes in each outcome variable (home price index and construction, financial, and other (total private excluding construction and finance) employment) during the boom (2003-2006) and bust (2007-2010) periods explained by investor activity, as defined in the text (see equation 2.4). The lower-bound estimates are obtained when controlling for 2006 household debt to income.

specification (equation 2.1), and second, the lower-bound estimates from the specification which controls for cross-sectional variation in 2006 household debt to income. The latter are lower-bound estimates, since variation in 2006 debt to income would control for variation in the extent to which locals engaged in local property investment. Chinco and Mayer (2016) find that local investors accounted for about about two-thirds of second-home buying over 2000-2007, though local prices were more sensitive to investments by out-of-towners.

Investor activity can explain a sizable fraction of the observed changes in economic activity in the boom and bust, particularly for construction and financial employment. The surge in property investment would stimulate new construction as well as demand for financial intermediaries such as real estate agents. In the baseline specification, investor activity can explain between 37 and 21 percent of the variation in construction and financial employment in the boom years (2003-2006), and 30 and 16 percent of the variation in the bust years (2007-2010).

The variation in other employment (total private minus construction and finance) explained by investor activity is much smaller, particularly in the boom (about 7 percent). This suggests that, at least in the boom, the employment gains generated by the surge in property investment were largely concentrated in construction and finance. In the bust, property investment can explain a larger share of the job losses in other employment – between 16 and 36 percent. The picture is similar for home prices, with investor activity accounting for a modest fraction of the change in home prices in the run-up, but a larger fraction in the bust, between 14 (lower-end) and 36 (baseline) percent.

For other employment, which accounts for about 85 percent of total employment, the losses associated with the property speculation in the bust years were larger than the gains in the boom. The asymmetry is consistent with the investment overhang hypothesis. That is consistent with the quantitative housing model of Boldrin et al. (2016) in which to irreversibility constraints on housing structures help explain employment losses in the bust. Housing structures cannot be put to use in other sectors where the marginal productivity of land is higher. In Rognlie, Shleifer, and Simsek (Forthcoming) misallocations are caused by the Zero Lower Bound, which places a cap on nonresidential investment and consumption.

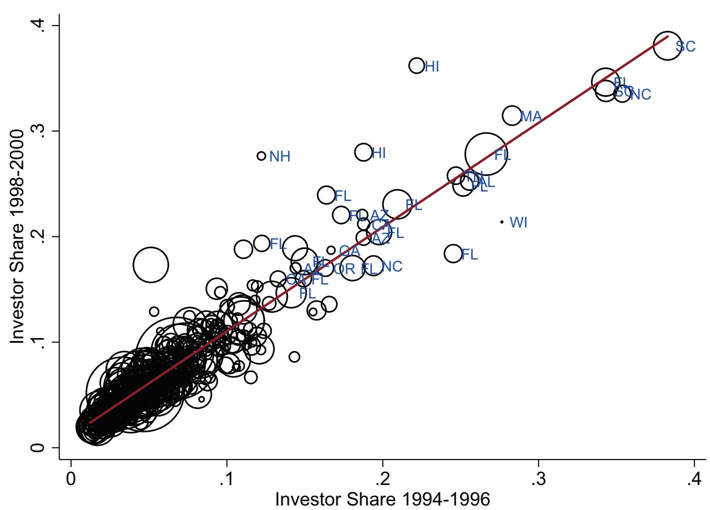
2.4.2 Robustness

The identifying assumption is that the 1998-2000 investor shares are uncorrelated with unobserved characteristics of counties affecting boom-bust dynamics.

In particular, it is possible that the investor shares measured over 1998-2000 are partly driven by investors' expectations about future home appreciation. The 1998-2000 period was chosen as a pre-boom period to avoid this concern, but it is plausible that it is not early enough.

To check against those concerns, I proceed in two ways. First, the investor share measured over an earlier period (1994-1996) is highly correlated (coefficient = 0.95) with the investor share measured over 1998-2000 (Figure 2.7). This ameliorates concerns about the cross-sectional variation in investor shares being driven by expected home appreciation. The correlation coefficient is 0.84 between the investor shares measured over 1994-1996 and 2004-2006 (not shown).

Figure 2.7: Stable Classification of Investor Counties



Source: HMDA. Figure plots investor shares in 1998-2000 against investor shares in 1994-1996. Observations weighted by the number of home sales in 2005-2006.

Instead, the high correlation is likely driven by the fixed appealing physical qualities of localities. To see that more directly, I isolate variation in the investor shares that is purely explained by observable physical characteristics, from the Natural Amenities Dataset. Specifically, I run the following 'first-stage' regression where P_i consists of the six variables in the Natural Amenities Dataset, including their squared terms, and the investor share is measured over 1998-2000 as in the rest of the paper. The physical characteristics of localities include measures of temperature, sunlight, topography, and water area.

$$Investor\ Share_i = \gamma P_i + v_i$$

$$\widehat{Investor\ Share}_i = \hat{\gamma} P_i$$

The correlation coefficient between these fitted values and the actual investor shares measured over 1998-2000 and 2004-2006 are 0.43 and 0.49, respectively. I then use the fitted values to repeat the main analysis of this paper.

$$Y_{it} = \alpha_i + \tau_t + \beta_t(\widehat{Investor\ Share}_i \times \tau_t) + \phi_t(Z_i \times \tau_t) + \epsilon_{it} \tag{2.5}$$

When using the variation in the fitted values explained by the observed physical characteristics of the localities, the results are very similar, as shown in Table 2.6. Investor counties grew faster in the boom years, and crashed harder. The coefficients are very similar to those in Table 2.3, after rescaling the fitted values to match the standard deviation in the investor shares. More specifically, using these estimates in the baseline case, variation in investor shares could explain 32, 16, 27, and 17 percent of the run-up in construction employment, home prices, financial employment, and other employment in the boom, and 13, 18, 10, and 11 percent of the decline, respectively, over 2007-2010.

In the baseline specification, the dependent variables are in log levels and do not include lags of the dependent variable, to allow for investor activity having potentially permanent effects. The benefit of estimating the models in first differences is that doing so is robust to counties having different time trends. When in first differences, the results are qualitatively very similar. Table 2.7 shows the coefficients associated with the interaction of investor shares and year dummies for the model in first differences (growth rates), which are very similar to the model in level including a lag of the dependent variable (not shown).

Economic activity associated with having higher investor shares peaked in 2005, the peak year of the boom. Growth in construction employment, home prices, and financial employment associated with investor activity was highest in 2005. Mortgage originations are a flow, so the peak in new economic activity in 2005 is reflected as a higher growth the year before (between 2004 and 2005). The effect of investor activity plateaued in the following year, with growth rates in 2006 not statistically different from zero. Over the next years, investor activity is associated with declining economic activity. The peak of the decline is in 2008 – in that year, counties with a 10 percent higher investor share (over 1998-2000) experienced on average 4.3, 6.3, 4.5, and 1.5 percent lower growth rates on average in home prices, construction employment, financial employment, and other employment. The negative association between growth rates continues until 2012, when once again investor activity is no longer significantly associated with differences in growth rates across counties.

Another interesting question is to what extent the variation in investor shares is distinct from the housing supply elasticity of Saiz (2010); counties with inelastic supply experienced a more pronounced boom and bust in economic activity

Mian and Sufi (2009). The answer is that they are quite distinct – the R-squared in a population-weighted (unweighted) regression of the elasticity against the investor share is 0.002 (0.004). The elasticity measure is available for a lower sample of counties, so it was not included in the main analysis though the results are essentially identical when including it. As another robustness check, I note that repeating the analysis of this paper using the vacation share from the 2000 Census instead of the investor share produces very similar results. This helps address potential concerns about measurement error based on reported owner-occupancy in HMDA, though as noted before, those concerns are ameliorated by the high correlation across years in the HMDA measure, as well as with related variables obtained from independent datasets.

Table 2.6: Investor Activity and Boom-Bust Dynamics – Using Variation in Investor Shares Explained by Natural Amenities

	Dependent variables:				
	$Originations_{it}$	$Home\ Prices_{it}$	$Constr.\ Emp_{it}$	$Fin.\ Emp_{it}$	$Other\ Emp_{it}$
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
β_{2003}	0.17***	0.03	0.08	-0.00	-0.02
	(0.05)	(0.08)	(0.05)	(0.05)	(0.02)
β_{2004}	0.34***	0.13	0.19***	0.05	0.01
	(0.12)	(0.09)	(0.06)	(0.06)	(0.02)
β_{2005}	0.63***	0.51***	0.41***	0.12**	0.06**
	(0.15)	(0.06)	(0.07)	(0.06)	(0.02)
β_{2006}	0.46***	0.77***	0.65***	0.18***	0.11***
	(0.14)	(0.08)	(0.08)	(0.07)	(0.03)
β_{2007}	-0.10	0.73***	0.61***	0.27***	0.13***
	(0.15)	(0.09)	(0.07)	(0.07)	(0.03)
β_{2008}	-0.74***	0.53***	0.38***	0.27***	0.11***
	(0.16)	(0.09)	(0.07)	(0.07)	(0.03)
β_{2009}	-0.52***	0.36***	0.17	0.20***	0.07**
	(0.17)	(0.10)	(0.13)	(0.07)	(0.03)
β_{2010}	-0.41**	0.17	-0.05	0.23***	0.05
	(0.17)	(0.16)	(0.13)	(0.08)	(0.03)
β_{2011}	-0.24	0.01	-0.14	0.26***	0.06*
	(0.19)	(0.17)	(0.12)	(0.08)	(0.04)
β_{2012}	-0.21	0.05	-0.06	0.27***	0.09**
	(0.13)	(0.13)	(0.11)	(0.08)	(0.04)
County, Year FE	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes
Within R-squared	0.80	0.84	0.59	0.35	0.56
# Observations	17136	10280	17710	17952	17930
# Counties	816	514	805	816	815

The table reports estimates based on the model

$$Y_{it}^{j} = \alpha_i + \tau_t + \beta_t(\widehat{Investor\ Share}_i \times \tau_t) + \phi_t(Z_i \times \tau_t) + \epsilon_{it}$$

where β_t is the coefficient associated with the interaction of the investor share over 1998-2000 and a year t dummy variable. $\widehat{Investor\ Share}_i$ denotes the fitted values from regressing the investor shares on the physical characteristics of localities from the Natural Amenities Scale (equation 2.2). Regressions include county fixed effects and time fixed effects for the sample of largest counties for which data is available. Additional controls include the interaction of year dummies with county characteristics Z_i: 2000 median home values, household income, the fraction of the population that is college-educated, poor, have an open mortgage, identify as white, are 55 years old or older, live in an urban area, and the manufacturing share of employment in 2000 from the QCEW. Standard errors clustered at the county-level.

Table 2.7: Investor Activity and Boom-Bust Dynamics – Models in First Differences

	Dependent variables:				
	$\Delta Originations_{it}$	$\Delta Home\ Prices_{it}$	$\Delta Constr.\ Emp_{it}$	$\Delta Fin.\ Emp_{it}$	$\Delta Other\ Emp_{it}$
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
β_{2003}	-0.02	-0.04	0.09	-0.07	0.01
	(0.07)	(0.06)	(0.06)	(0.06)	(0.02)
β_{2004}	0.14	-0.03	0.15***	0.04	-0.00
	(0.10)	(0.06)	(0.05)	(0.05)	(0.02)
β_{2005}	0.05	0.34***	0.31***	0.07	0.01
	(0.11)	(0.11)	(0.07)	(0.05)	(0.02)
β_{2006}	-0.83***	0.04	0.06	0.02	0.01
	(0.12)	(0.07)	(0.06)	(0.04)	(0.02)
β_{2007}	-0.95***	-0.33***	-0.25***	-0.01	-0.06***
	(0.18)	(0.08)	(0.08)	(0.04)	(0.02)
β_{2008}	-0.87***	-0.39***	-0.64***	-0.14***	-0.11***
	(0.14)	(0.11)	(0.09)	(0.04)	(0.02)
β_{2009}	-0.32**	-0.44***	-0.61***	-0.19***	-0.14***
	(0.12)	(0.10)	(0.09)	(0.04)	(0.02)
β_{2010}	-0.17**	-0.37***	-0.27***	-0.02	-0.05***
	(0.08)	(0.11)	(0.07)	(0.04)	(0.02)
β_{2011}	0.00	-0.24***	-0.22***	-0.07**	-0.01
	(0.07)	(0.09)	(0.06)	(0.03)	(0.01)
β_{2012}	-0.13*	0.03	-0.04	-0.05	-0.00
	(0.08)	(0.05)	(0.06)	(0.04)	(0.01)
County, Year FE	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes
Within R-squared	0.71	0.74	0.47	0.18	0.48
# Observations	16460	9804	17052	17283	17262
# Counties	823	516	812	823	822

The table reports estimates based on the model

$$\Delta Y_{it}^{j} = \alpha_i + \tau_t + \beta_t(Investor\ Share_i \times \tau_t) + \phi_t(Z_i \times \tau_t) + \epsilon_{it}$$

where β_t is the coefficient associated with the interaction of the investor share over 1998-2000 and a year t dummy variable. The dependent variable here is in first-differences of the log level of the dependent variable. Regressions include county fixed effects and time fixed effects for the sample of largest counties for which data is available. Additional controls include the interaction of year dummies with county characteristics Z_i: 2000 median home values, household income, the fraction of the population that is college-educated, poor, have an open mortgage, identify as white, are 55 years old or older, live in an urban area, and the manufacturing share of employment in 2000 from the QCEW. Standard errors clustered at the county-level.

2.5 Conclusion

Much of the lively debate on the origins of the housing boom has focused on the extent to which households of different credit scores and income contributed to the run-up in debt in the early and mid 2000s (Mian and Sufi 2009; Adelino, Schoar, and Severino 2016; Albanesi, Giorgi, and Nosal 2017). This paper contributes to that literature on a relatively unexplored dimension of the housing boom – property investors, existing home-owners acquiring new property. This paper documents that investor counties experienced higher speculation (flip rates) in the peak boom years. They also experienced a pronounced boom-bust cycles in mortgage originations, home prices, and construction employment.

Understanding the role played by property investors might be important for determining the presence of a real estate bubble. From 2003-2006, when second and third-home buying flourished, the home ownership rate barely budged from 68.3 to 68.8 percent. During this time second-home purchases contributed more to housing debt than first-time home owners or equity extractors (Bhutta 2015). In this light, the peak boom years are better understood as a period of second-home buying and speculation. This paper argues that investor activity played an important role in determining economic activity over 2003-2010. This is in line with predictions of housing bubble models showing that relatively few optimists can influence economic activity (Piazzesi and Schneider 2009; Burnside, Eichenbaum, and Rebelo 2016; DeFusco, Nathanson, and Zwick 2017).

This paper shows that the new properties acquired were concentrated in counties with pleasant qualities such as warm winters and waterfronts. These counties experienced a surge in speculation and economic activity over 2003-2006, and a collapse in the following years. Property investment could explain about 30 percent of the variation in construction and financial employment over 2003-2010. Property investment is also associated with sizable losses in home prices and other employment categories over 2007-2010, of about a third of the observed declines.

Chapter 3

Can Risky Rural-Urban Migration Explain the Flow of Capital from Developing to Advanced Economies?

3.1 Introduction

According to the textbook neoclassical growth model, countries with faster productivity growth should attract more foreign capital. Instead, high growth countries tend to export capital. Gourinchas and Jeanne (2013) name this puzzling correlation the Allocation Puzzle, and document that the reason is a saving wedge – savings in high growth countries are much larger than what the textbook neoclassical growth model predicts. There are many potential explanations, including consumption habits (Carroll, Overland, and Weil 2000), borrowing constraints for both households and firms (Sandri 2014; Coeurdacier, Guibaud, and Jin 2015), and demographics such as a growing population of young savers (Modigliani 1970; Chamon, Liu, and Prasad 2013). Carroll and Jeanne (2015) suggest a different answer: countries with higher productivity growth might have higher uninsurable idiosyncratic risk. An open question is what plausible microfounded household income process could deliver both growth and net capital outflows.

This paper provides an answer by focusing on a specific type of risk – the migration of workers from rural to urban areas within a country. Many rapidly growing countries have urbanized quickly in the last decades. In China, for example, the urban population grew from 20 to 52 percent between 1980 and 2012. Urbanization generates income growth, as labor relocates to the higher productivity urban sector (Williamson 1988; Henderson 2003). However, migration also entails substantial idiosyncratic risk (Stark 1993; Jaeger et al. 2010). One driver of income volatility is that migration in many developing countries tends to be circular: workers move back and forth between rural and urban areas, and so income fluctuates based on location. Reviewing several country studies, Skeldon (2012) notes that circular migration is so prevalent "regular short-term movement back and forth between village and town far exceed any longer-term migration." In China, the internal passport system *hukou* explicitly encourages circular migration, by in some cases banning rural households from permanently locating to urban centers (Hare 1999; Zhao 2005).

This paper models the saving motives of residents in a country undergoing urbanization characterized by circular migration, using a heterogeneous agent model with three worker types – urban, migrant, and rural. Workers move to and fro sectors as determined by exogenous transition probabilities calibrated to match migration flows and the pace of urbanization in China. When workers are employed in the urban sector, they earn the higher urban wage, and save as self-insurance against location shocks. Rural workers, on the other hand, run down the savings they accumulated while employed in the city. As the country urbanizes, the population shifts towards the savers, and so aggregate savings rise. Income also grows because labor is more productive in the urban sector. Investment also rises for the same reason, though it does not match the increase in savings. As a result, the urbanization process results in both income growth as well as net capital outflows.

The model predicts that, all else equal, countries with a larger increase in the urban population (the savers) would be more likely to export capital.[1] That is indeed the case in the data. Countries with larger increases in the share of the urban population over 1980-2000 experienced on average lower net capital inflows relative to GDP (the correlation coefficient is -0.43).[2] This may shed light on why some fast-growing developing countries tend to export capital (e.g. China) but others (e.g. Chile) do not. The answer suggested in this paper is countries like Chile were already close to the steady state urban population, and so did not recently undergo a rapid and dramatic urbanization process the way China has. This contributes to the literature on the Allocation Puzzle. Models focusing on the upstream flow of capital, such as Quadrini, Mendoza, and Rios-Rull (2009); Caballero, Farhi, and Gourinchas (2008); Angeletos and Panousi (2011) can explain why capital flows from developing countries to advanced economies, but are silent on the cross-country correlation between growth and capital outflows within developing countries documented in Gourinchas and

[1]While the model is calibrated to match China's urbanization process, the key feature of the model – that urbanization is characterized by circular migration – is a feature of many countries; see for example Lucas (1993); Collinson et al. (2006); Deshingkar and Akter (2009); Newland (2009); Thom (2010).

[2]This is shown in Figure 3.5 for the sample of 68 developing countries in Gourinchas and Jeanne (2013).

Jeanne (2013).

The model also predicts a positive correlation across countries between economic growth and capital outflows, based on productivity differences. In countries with higher urban sector productivity, urbanization leads to both *i*) larger income gains from labor relocation, and *ii*) a larger urban-rural wage gap which increases income volatility from location shocks and leads to larger net capital outflows. I also explore the effects of reducing frictions prohibiting migrants from permanently relocating to the city. When those frictions are lowered, there is less income volatility associated with worker location shocks, and so savings decrease. This suggests that the 2014 *hukou* reforms in China, which relaxed legal barriers to migration, may be part of the reason why China's consumption share of GDP has increased since.

This paper is most closely related to the literature arguing that precautionary savings associated with economic growth explain why fast-growing developing countries tend to export capital, reviewed in Gourinchas and Rey (2013). Carroll and Jeanne (2015) argue the "growth-to-saving puzzle can be explained ... if the bargain that countries make when they embark on a path of rapid development involves not only a pickup in productivity growth but also an increase in the degree of idiosyncratic risk." This paper contributes to this literature by focusing on risks associated with migration. Other work emphasizes other aspects of economic growth such as financial development. In Sandri (2014) growth accelerations cause net capital exports because credit constraints imply entrepreneurs need to save more than they invest for precautionary motives. In Buera and Shin (2017) economic reform leads to productivity growth and higher savings, though investment lags because of domestic financial frictions. Similarly, in Song, Storesletten, and Zilibotti (2011) and Bacchetta and Benhima (2015) external surpluses are generated by the interaction between productivity growth and financial frictions.

3.2 Model

The economy is represented by a heterogenous agent model with aggregate certainty. There are three ex-ante identical household types – rural, migrant, and urban – who are subject to exogenous location shocks given by a first-order Markov chain. Both urban and migrant workers earn the urban wage, though migrants are much more likely to move to the rural sector. Rural workers earn the rural wage.

3.2.1 Firms

Goods in the urban and rural sectors are produced by competitive firms with Cobb-Douglas production functions. Total factor productivity (TFP) A_ℓ in each sector ℓ is assumed to be constant, with productivity higher in the urban sector, $A_u > A_r$. To make goods $Y_\ell(t)$ in sector ℓ at time t, urban firms employ migrant $M(t)$ and urban $U(t)$ workers, while rural firms employ only rural $R(t)$ workers. Firms also employ capital $K_u(t)$ in the urban sector and $K_r(t)$ in the rural. Production each period is given by:

$$Y_u(t) = A_u K_u(t)^\alpha (M(t) + U(t))^{1-\alpha} \tag{3.1}$$

$$Y_r(t) = A_r K_r(t)^\alpha R(t)^{1-\alpha} \tag{3.2}$$

where $\alpha \in (0, 1)$ is elasticity of output with respect to capital. Capital in both sectors depreciates at rate δ, and earns the world interest rate r^*. Output prices in both sectors are normalized to 1. Firms maximize profit with respect to their labor and capital demand. These assumptions, together with constant TFP in each sector, pin down the equilibrium capital-per-worker ratio in each sector, which is constant each time period:

$$\frac{K_u(t)}{M(t) + U(t)} = \left(\frac{A_u \alpha}{r + \delta}\right)^{1/(1-\alpha)} \tag{3.3}$$

$$\frac{K_r(t)}{R(t)} = \left(\frac{A_r \alpha}{r + \delta}\right)^{1/(1-\alpha)} \tag{3.4}$$

Similarly, equilibrium wages are also constant, since the capital-per-worker ratio is constant in each sector.

$$w_u = A_u(1 - \alpha)\left(\frac{K_u(t)}{U(t) + M(t)}\right)^{\alpha} \tag{3.5}$$

$$w_r = A_r(1 - \alpha)\left(\frac{K_r(t)}{R(t)}\right)^{\alpha} \tag{3.6}$$

3.2.2 Households

The economy consists of many infinitely lived individuals. In particular, there is a continuum of agents of total mass equal to one. Each household consists of one agent and I use the terms household, agent, or worker interchangeably.

Worker location, and thereby income, is given by an exogenous Markov process, denoted by

$$P = \begin{pmatrix} p_{uu} & p_{mu} & p_{ru} \\ p_{um} & p_{mm} & p_{rm} \\ 1 - p_{uu} - p_{um} & 1 - p_{mu} - p_{mm} & 1 - p_{ru} - p_{rm} \end{pmatrix} \tag{3.7}$$

For example, given that a worker is currently a migrant, she remains a migrant with probability p_{mm}, becomes an urban worker with probability p_{mu}, or returns to the rural sector with probability $1 - p_{mm} - p_{mu}$. To focus only on precautionary savings associated with circular migration, labor supply is perfectly inelastic. There is therefore no unemployment in either of the sectors.

Households are ex-ante identical, but may differ in their asset holdings s or their employment location ℓ. Given values of $\{w_u, w_r, r^*\}$ and initial values for asset holdings and location $\{s_0, \ell_0\}$, the household chooses a policy for $\{s_{t+1}\}_{t=0}^{\infty}$ to maximize

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(c_t) \tag{3.8}$$

subject to

$$s_{t+1} = \begin{cases} (1 + r^*)s_t + w_u - c_t & \text{if } \ell = U \text{ or } \ell = M \\ (1 + r^*)s_t + w_r - c_t & \text{if } \ell = R \end{cases}$$

and $s_{t+1} \in \mathcal{S}$ and $U(c) = \frac{c^{1-\rho}}{1-\rho}$, where ρ is the coefficient of relative risk aversion.[3] Workers earns the urban wage w_u when they are a migrant ($\ell = M$) or urban worker ($\ell = U$), and the rural wage otherwise.

Let $V(s, \ell)$ be the value of the objective function of a household with asset holdings s and location ℓ. $V(s, \ell)$ is defined as the solution to the dynamic program:

[3] I discretize the asset space $\mathcal{S}$ into 251 equally spaced assets in the grid $[-2, 2500]$, and use a cubic spline to interpolate between gridpoints.

$$V(s,\ell) = \max_{s'} \; U((1+r^*)s + w_\ell - s') + \beta \mathbb{E}_t[V(s',\ell')|\ell)]$$

where s' are next period's assets and the policy function $s' = g(s,\ell)$ maps the current period's (s,ℓ) pair into an optimal choice of assets to carry into the next period.

Define the unconditional distribution of (s,ℓ) pairs, $\lambda_t(s,\ell) = Prob(s_t = s, \ell_t = \ell)$. The policy function $g(s,\ell)$ together with the transition matrix P for location induce a law of motion for the distribution λ_t:

$$\lambda_{t+1}(s',\ell') = \sum_{\ell} \sum_{s:s'=g(s,\ell)} \lambda_t(s,\ell) p_{\ell\ell'} \tag{3.9}$$

At any given moment, aggregate wealth in the economy (S for the stock of saving) equals the sum of individual asset holdings, where the sum is weighted by the mass of households holding the same assets in each location.

$$S_t = \sum_{\ell} \sum_{s} \lambda_t(s,\ell) s$$

Assets can be hold domestically as capital ($K(t) = K_u(t) + K_r(t)$), or as net foreign assets (NFA_t), and so therefore

$$NFA_t = S_t - K_t$$

The key assumption here is that domestic and foreign assets are perfectly substitutable because of the small country assumption that both assets pay the world interest rate r^*.

3.2.3 Stationary Equilibrium

Let $\bar{U}$, $\bar{M}$ and $\bar{R}$ denote the steady state population of consumers located in the urban, migrant and rural locations, respectively. The steady state populations are determined by the Markov process P.

A stationary equilibrium consists of a policy function $g(s,\ell)$, a probability distribution $\lambda(s,\ell)$, and positive reals $(\bar{U}, \bar{M}, \bar{R}, K, w^u, w^r, r^*, C, S, NFA)$ such that:

1. $\begin{pmatrix} \bar{U} \\ \bar{M} \\ \bar{R} \end{pmatrix} = \begin{pmatrix} p_{uu} & p_{mu} & p_{ru} \\ p_{um} & p_{mm} & p_{rm} \\ 1 - p_{uu} - p_{um} & 1 - p_{mu} - p_{mm} & 1 - p_{ru} - p_{rm} \end{pmatrix} \begin{pmatrix} \bar{U} \\ \bar{M} \\ \bar{R} \end{pmatrix}$

2. $\bar{U} + \bar{M} + \bar{R} = 1$

3. The prices (w_u, w_r, r^*) satisfy first order conditions for firms in equations (3.3)-(3.6).

4. The policy function $g(s,\ell)$ solves the household's optimization problem as stated in equation (3.8).

5. The probability distribution $\lambda(s,\ell)$ is a stationary distribution associated with P and $g(s,\ell)$. That is, it satisfies

$$\lambda(s',\ell') = \sum_{\ell} \sum_{s:s'=g(s,\ell)} \lambda(s,\ell) p_{\ell\ell'}$$

6. Aggregate consumption is $C = \sum_{\ell} \sum_{s} \lambda(s,\ell) c(s,\ell)$ where $c(s,\ell)$ is the policy rule for consumption that is implied by the policy rule for next-period wealth $g(s,\ell)$. Similarly, aggregate wealth is $S = \sum_{\ell} \sum_{s} \lambda(s,\ell) s$ and $S = NFA + K$ where NFA is net foreign assets.

7. The aggregate resource constraint holds: $C + S = (1+r)S + w^u(\bar{U} + \bar{M}) + w^r \bar{R}$

3.3 Calibration

The transition probability matrix P is calibrated using data on migration flows and the pace of urbanization in China. Six parameters need to be calibrated between the steady state population shares of rural, urban, and migrant workers H, the transition probability matrix P, and the steady state relation between the two:

$$H := \begin{pmatrix} \bar{U} \\ \bar{M} \\ 1 - \bar{M} - \bar{R} \end{pmatrix}$$

$$P = \begin{pmatrix} p_{uu} & p_{mu} & p_{ru} \\ p_{um} & p_{mm} & p_{rm} \\ 1 - p_{uu} - p_{um} & 1 - p_{mu} - p_{mm} & 1 - p_{ru} - p_{rm} \end{pmatrix}$$

$$H = PH$$

The model is calibrated using:

1. $p_{uu} = 0.99$. Permanence in the urban sector is very high. Some leakages are needed, since otherwise all households would be urban workers in steady state. Leakages from the urban sector could occur because of marriage, injuries, natural disasters, preference shocks, etc which can encourage workers to return to rural areas.

2. $p_{ur} = 0$: Workers in the urban location cannot directly move to the rural sector, they first have to become migrants.

3. $p_{ru} = 0$: Workers in the rural location cannot directly move to the urban sector, they first have to become migrants.

4. $\bar{U} = 0.7$: In steady state the population is largely urban. Similar values to $\bar{U}$ yield the same qualitative conclusions.

5. $\bar{M} = 0.1$: Based on data from the official 2005 Population Sample Survey of China, Taylor (2011) finds that 147 million Chinese are migrants, or about 10 percent of the population. In equilibrium I set $\bar{M} > 0$ due to legal frictions in China's registration system hukou that make it difficult for workers to stay permanently in the city.[4]

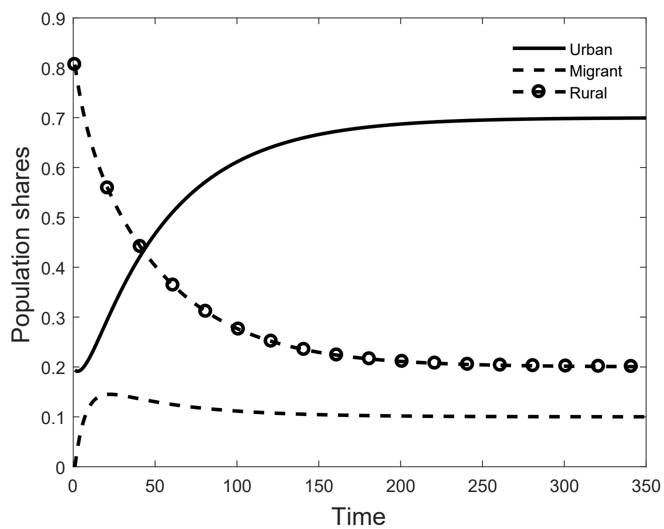
6. $p_{rr} = 0.9727$: This is chosen to match the average growth rate of the rural population in the model with that of the World Bank's data on urbanization in China between 1980 and 2012. The initial values used are $U(1980) = .19$, $M(1980) = 0$, and $R(1980) = 0.81$. The rural population on average declined by 1.62 percent each year.

In sum, the transition probability matrix and steady state populations used in the baseline calibration are:

$$P = \begin{pmatrix} p_{uu} = .9900 & p_{mu} = .0700 & p_{ru} = .0000 \\ p_{um} = .0100 & p_{mm} = 0.8755 & p_{rm} = .0273 \\ p_{ur} = .0000 & p_{mr} = .0545 & p_{rr} = .9727 \end{pmatrix}$$

[4] Individuals with rural hukou lack access to a wide range of urban services: they cannot enroll their children in urban public schools (or have to pay high fees to do so), and have less access to public health care, legal assistance, the pension system, and social services (Zhan 2005; Wing Chan and Buckingham 2008). See Zhao (2005) for a review on hukou and temporary migration in China.

Figure 3.1: Transition to an urban economy



and

$$H = \begin{pmatrix} \bar{U} = .7 \\ \bar{M} = .1 \\ \bar{R} = .2 \end{pmatrix}$$

Figure 3.1 shows the simulated path of urban, rural and migrant workers, compared to the actual data (from the World Bank), assuming that the number of migrants in 1980 is zero. The economy is at first largely rural, but grows more urban with time. Eventually it reaches its steady state distribution where 70 percent of households are urban.

Another key parameter is the size of the urban-rural wage gap. The average gap in China was about 2.5 between 1980 and 2012 according to the National Bureau of Statistics in China. This wage gap pins down the urban-rural TFP ratio. The coefficient of relative risk aversion is $\rho = 2$. The annual world interest rate is $r^* = .04$. $\beta = 0.957$ sets the steady state ratio of wealth to income to 5.7. This ratio is in line with the wealth to income ratios estimated in Piketty and Zucman (2013) for a few advanced economies (4-6 for USA, UK, Germany and France).

The economy's initial conditions also need to be determined. The initial state is taken from China's conditions in 1980. Then, the economy was largely rural and there was little to no migration due to legal bans prior to Deng Xiaoping's reforms. The rural share of the population in 1980 was 81% and the urban 19%.

$$U(0) = 0.19; \quad M(0) = 0.00 \quad R(0) = 0.81$$

Because of the tight controls on the rural population, I assume that at time $t = 0$ rural workers did not own any capital ($K_r(0) = 0$). Urban workers, on the other hand, held the level of capital ($K_u(0)$) consistent with the parameters in the model given the production process

$$K_u(0) = U_0 \left(\frac{A_u \alpha}{1/\beta - 1 + \delta} \right)^{\frac{1}{1-\alpha}} = 29.62$$

Other similar initial values, such as allowing rural workers to have the same level of initial capital, yield the same qualitative results.

3.4 Transition to an urban economy

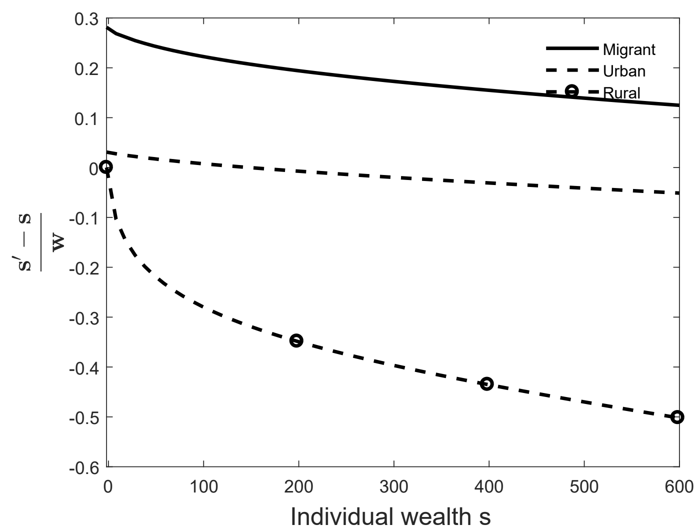
Starting off from a largely rural population share, the economy then advances to its steady state where the urban share of the population is 70 percent. As the economy urbanizes, saving increases as the population shifts towards the savers. Income also grows because of the relocation of workers to the more productive urban sector. The capital stock also rises, though less so than saving. As a result, the urbanization process leads to income growth and net capital outflows.

When workers earn the higher urban wage, they have an incentive to save to insure against the risk of returning to the low-wage rural sector. I define the saving rate as the change in wealth implied by the optimal policy divided by current period labor income: $\frac{a'-a}{w}$. Figure 3.2 plots saving rates by worker type and assets. Migrants have the highest saving rate. They earn the urban wage, and face a 5.45 percent chance of moving to the rural sector in the next period, where permanence is quite high (97.27 percent chance of staying there from one period to the next). In the baseline calibration, the saving rate for migrants ranges from 14 to 28 percent, when they are asset poor.

Urban workers have the next highest saving rates. About 20 percent of the urban workers (those with a wealth-to-income ratio lower than 4.14) have a positive saving rate. The rest have negative saving rates, though only moderately so. When the ratio of wealth to income drops below 4.14, urban workers begin to save again, and as a result all urban workers have a positive stock of saving. Rural workers, on the other hand, run down the saving they accumulated while earning in the city. They have negative saving rates for all possible asset levels (Figure 3.2).

Figure 3.3 plots the stationary distribution of wealth λ by worker type, that is, for any feasible pairs (s', ℓ'), $\lambda(s', \ell') = Prob(s = s', \ell = \ell')$. The figure shows all urban and migrant workers have a positive stock of savings. The stationary wealth to income ratio for the average migrant is 5.74 and 5.51 for the average urban worker. This is the saving stock migrants then use to smooth consumption when in the rural sector. In equilibrium, close to 20 percent of rural workers (4 percent of total population) have fully run down their wealth and are at the model's borrowing constraint, as shown in Figure 3.3.

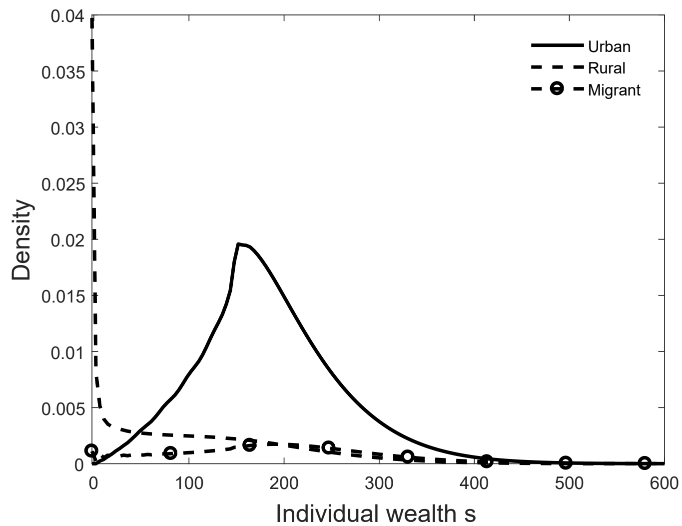
Figure 3.2: Saving rates of workers



As the economy urbanizes, the population shifts towards the agents with the higher saving rates in the model. As a result, the stock of saving grows, as do net foreign assets. Figure 3.4 plots the time path of the total stock of saving S, domestic capital K, and net foreign assets NFA. The net foreign assets equal the difference between total saving and domestic capital. With the urbanization process, all three series increase. Capital increases because of the flow of labor into the urban sector, which is more productive. This raises aggregate productivity and tends to raise the marginal product of capital. Capital therefore increases to ensure the marginal product of capital is equal to the world interest rate net of depreciation. Similarly, income rises as well.

The net foreign asset position also increases in the long-run because the increase in the stock of saving outpaces the increase in capital. This is true long

Figure 3.3: Invariant density function of wealth



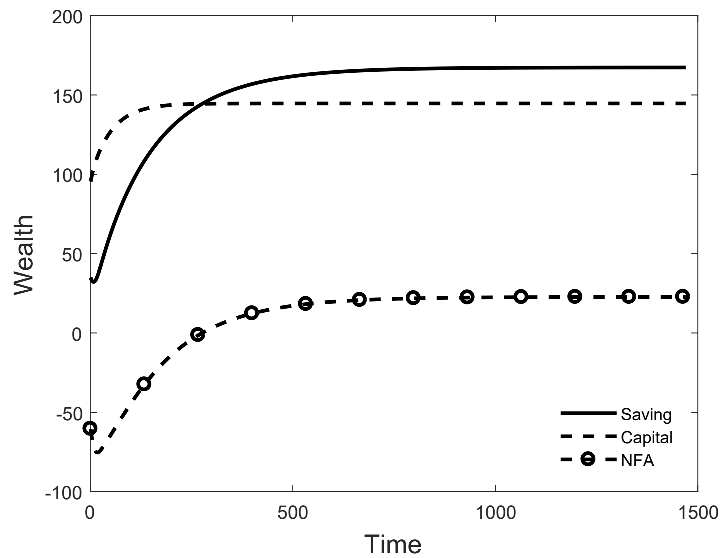
term, but in the first periods net foreign assets slightly decrease. The reason is that capital investment quickly rises as workers pour into the urban sector. In contrast, it takes somewhat longer for the agents in the model to build their stock of saving. After a few periods, however, the increase in net foreign assets quickly outpaces that of capital. The model thus predicts that as the economy urbanizes, income grows, as do the net foreign asset position of the country.

3.5 Applications

The model predicts that, all else equal, countries with a larger increase in the urban population (the savers) would be more likely to export capital.[5] To see this, suppose there are two identical economies, so both have the same steady state level of net foreign assets. Suppose one is already at (or is close to) its

[5]While the model is calibrated to match China's urbanization process, the key feature of the model – that urbanization is characterized by circular migration – is a feature of many countries; see for example Lucas (1993); Collinson et al. (2006); Deshingkar and Akter (2009); Newland (2009).

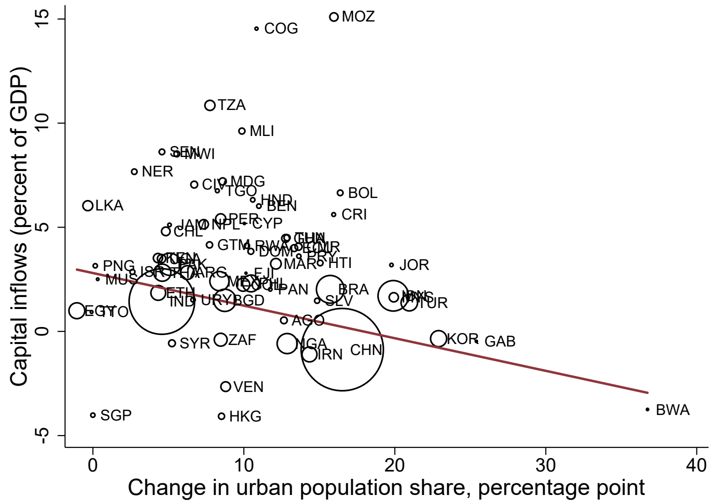
Figure 3.4: Aggregate wealth dynamics towards stationary equilibrium



stationary equilibrium; then by definition, neither savings nor the urban share of the population would change (or increase much). On the other hand, an economy with a largely rural starting point, such as the one examined in the previous section, would be on its way to experiencing a large increase in the urban population as well as an increase in net foreign assets.

There is some evidence in the data that countries who have urbanized the most in recent decades have also seen larger improvements in their net foreign asset position. Figure 3.5 plots the change in the fraction of the urban population between 1980 and 2000 (horizontal axis) against average net capital inflows relative to GDP, for the sample of countries in Gourinchas and Jeanne (2013). There is a negative correlation ($\rho = -0.43$), which means that countries with only small increases in the urban population tend to have larger net capital inflows.

Figure 3.5: Countries with larger increases in the urban population tend to have lower net capital inflows

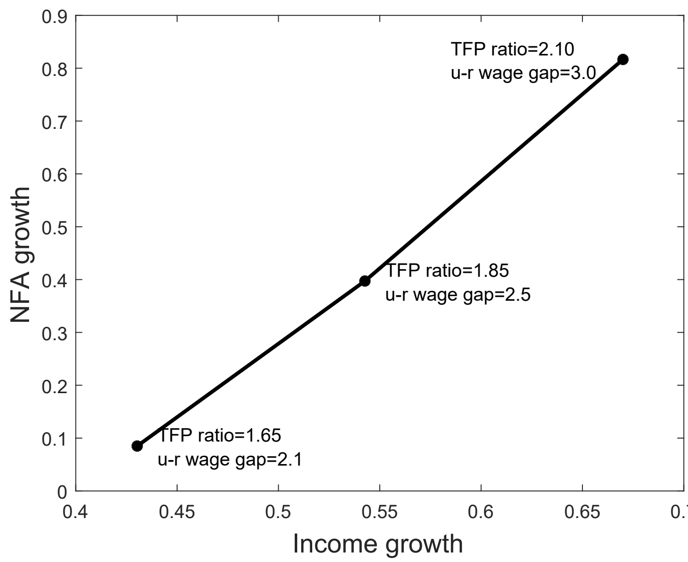


Changes in the urban population share and average capital inflows between 1998 and 2000. The sample of 68 non-OECD countries and capital inflow data are obtained from Gourinchas and Jeanne (2013). Urban population shares obtained from the World Bank Development Indicators.

This may shed light on why some fast-growing developing countries tend to export capital (e.g. China) but others (e.g. Chile) do not. The answer suggested in this paper is that countries like Chile were already close to its steady state urban population, and so did not undergo the rapid and dramatic urbanization process the way China did in recent decades. This contributes to the literature on the Allocation Puzzle. Models focusing on the upstream flow of capital, such as Quadrini, Mendoza, and Rios-Rull (2009); Caballero, Farhi, and Gourinchas (2008); Angeletos and Panousi (2011) can explain why capital flows from developing countries to advanced economies, but have more trouble explaining the Allocation Puzzle within developing nations.

The model can also predict part of the cross-sectional correlation between economic growth and capital outflows, based on differences in urban TFP (A_u in

Figure 3.6: In economies with higher TFP, income and NFA grow more



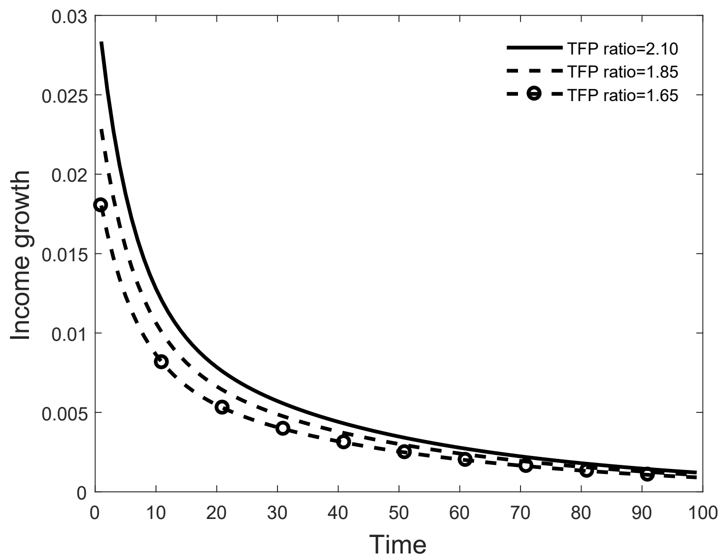
The figure plots growth in net foreign assets (y-axis) against income growth (x-axis) in economies with different urban-rural TFP ratios $\frac{A_u}{A_r}$ that are otherwise identical. Growth is over the first 100 periods of the simulation, e.g. income growth $= \frac{Y(100)}{Y(1)} - 1$.

equation (3.3)). This is summarized in Figure 3.7, which plots growth in countries' net foreign assets (y-axis) in the transition towards equilibrium, against income growth resulting from productivity gains due to urbanization (x-axis). The model predicts that, all else equal, countries that experience faster growth also export more capital.

The reason is that urbanization leads to higher economic growth in countries with higher urban TFP, because of the larger gains associated with labor relocation. This is shown in Figure 3.7, which plots income growth (y-axis) against time (x-axis) in three identical economies that differ only in the level of urban TFP.

Countries with larger TFP also experience larger increases in net foreign

Figure 3.7: Income growth, by TFP ratio

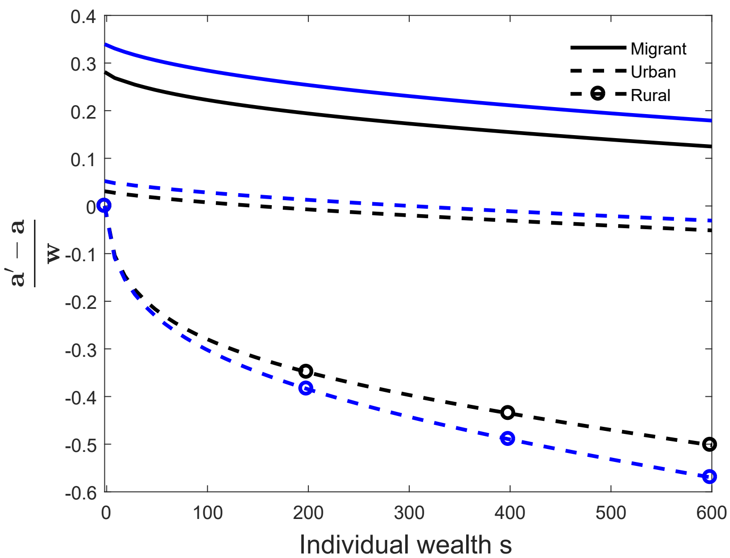


The figure shows income growth in economies with different urban-rural TFP rations $\frac{A_u}{A_r}$ that are otherwise identical. Income growth is period-to-period $\frac{Y(t)}{Y(t-1}$-1.

assets. The reason is that, all else equal, higher urban TFP generates a larger urban-rural wage gap, thereby raising the volatility of income across locations. Therefore, urban and migrant saving rates are higher in economies with higher TFP. This is shown in Figure 3.8, which plots saving rates (y-axis) by worker types and assets for two different economies, which only differ in the level of urban TFP. As the economy urbanizes, aggregate saving increases more in the economy where the savers have a higher saving rate.

I also examine the effects of removing barriers for the permanent relocation of workers in the city. In particular, I compare stationary equilibria with different values for p_{mu}, the probability with which migrant workers become urban workers in the next period. All else equal, higher chances of becoming an urban worker imply lower chances of returning to the rural sector. The lower risk

Figure 3.8: Saving rates of workers, by TFP ratio



The black curves correspond to saving rates for the baseline economy with TFP ratio = 1.85. Blue lines above are for the economy with TFP ratio = 2.10

induces migrants to save less. As a result, larger p_{mu} leads to a lower steady state ratio of net foreign assets to income.

The motivation for this exercise is the recent relaxation of hukou legal barriers to urban migration. Historically, households with rural hukou have faced various disadvantages in the urban sector, such as difficulties buying property, enrolling children in schools, lower access to health care, etc (Zhan 2005; Wing Chan and Buckingham 2008). Since 2014, however, some hukou restrictions have been lifted.[6]

In the baseline calibration, $p_{mu} = .07$. I now let it range from .06 to .08, while continuing with the restrictions outlined in the baseline calibration that: $p_{ru} = p_{ur} = 0$, $p_{uu} = 0.99$, $\bar{M} = .1$, and p_{rr} is chosen so that the average

[6]See, for example: https://www.theguardian.com/world/2014/jul/31/china-reform-hukou-migrant-workers

growth rate of the rural population in the model matches that of the data, from 1980-2012. Note that since

$$\bar{U} = p_{uu}\bar{U} + p_{mu}\bar{M} + p_{ru}\bar{R}$$

$$\bar{U} = 0.99\bar{U} + p_{mu}0.1$$

$$\bar{U} = 10p_{mu}$$

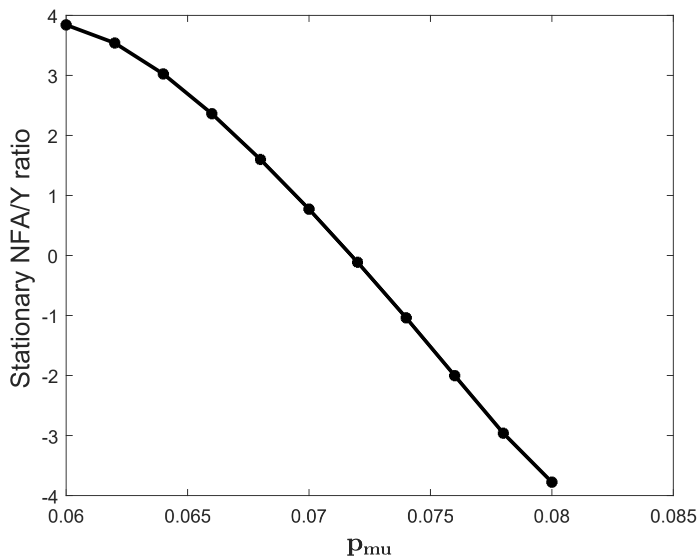
letting p_{mu} range from .06 to .08 is equivalent to shifting the steady state urban population from .6 to .8.

Lowering the barriers to urbanization reduces income volatility and so diminishes household demand for precautionary saving. As a result, the equilibrium ratio of net foreign assets to income falls when p_{mu} is higher. Figure 3.9 plots the steady state ratio of net foreign assets to income (y-axis) for economies with different p_{mu} but who are otherwise identical. The ratio of net assets to income falls when p_{mu} is higher. In recent years, the consumption share of GDP in China has risen. The model suggests a partial explanation might be the reduction in the precautionary saving motive associated with circular migration.

3.6 Conclusion

There are many explanations of why fast-growing countries tend to export capital. One of the leading explanations is that some aspect of the growth experience leads to higher idiosyncratic uninsurable risk, which translates to high savings and high demand for safe foreign assets. What are some of those specific risks? And why do they matter for macroeconomics?

Figure 3.9: Lowering barriers to urbanization (higher p_{mu}) lowers NFA/Y



The figure plots stationary net foreign assets to income (y-axis) against p_{mu} (x-axis) which is the probability that a migrant becomes an urban worker in the next period. The economies are otherwise identical.

The contribution of this paper is to provide an answer based on household migration from rural to urban areas. Many fast-growing countries have urbanized quickly in recent decades. In China, for example, between 300 and 400 million people migrated to cities between 1979 and 2009 (Chan 2013). This paper posits that risks associated with urbanization, driven by circular migration, could explain part of the reason why high growth countries tend to save in excess of investment.

I model a heterogeneous agent economy with three worker types – rural, migrant, and urban – who experience location shocks calibrated to match migration flows and the pace of urbanization in China. Workers in the high wage urban sector accumulate savings as self-insurance against location shocks. As the economy urbanizes, savings grow faster than investment, and so the country exports capital. At the same time, income grows as workers relocate to the higher productivity urban sector. The model thus generates the within-country stylized fact that growth episodes are accompanied by increases in saving and net foreign assets.

The model also generates cross-sectional predictions supported in the data. One prediction is that countries with larger increases in the share of the urban population are more likely to be capital exporters. The model also predicts a positive correlation between income growth (driven by urbanization) and net capital outflows. Lastly, the model suggests that removing barriers to urbanization diminishes the need for saving, and could therefore partly explain why China's consumption share of GDP has risen recently.

Bibliography

Adelino, Manuel, Antoinette Schoar, and Felipe Severino. 2012. "Credit Supply and House Prices: Evidence from Mortgage Market Segmentation." NBER Working Papers 17832, National Bureau of Economic Research, Inc. URL http://EconPapers.repec.org/RePEc:nbr:nberwo:17832.

———. 2016. "Editor's Choice Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class." *Review of Financial Studies* 29 (7):1635–1670. URL https://ideas.repec.org/a/oup/rfinst/v29y2016i7p1635-1670..html.

Albanesi, Stefania, Giacomo De Giorgi, and Jaromir Nosal. 2017. "Credit Growth and the Financial Crisis: A New Narrative." Unpublished working paper.

Alexandrov, Alexei and Sergei Koulayev. 2017. "No Shopping in the U.S. Mortgage Market: Direct and Strategic Effects of Providing Information." Unpublished working paper 2017-01, CFPB Working Paper.

Amiti, Mary and David E. Weinstein. 2018. "How Much Do Idiosyncratic Bank Shocks Affect Investment? Evidence from Matched Bank-Firm Loan Data."

Journal of Political Economy 126 (2):525–587. URL https://doi.org/10.1086/696272.

Anenberg, Elliot, Aurel Hizmo, Edward Kung, and Raven Molloy. 2016. "The Effect of Mortgage Credit Availability on House Prices and Construction: Evidence from a Frontier Estimation Approach." Working paper.

Angeletos, George-Marios and Vasia Panousi. 2011. "Financial integration, entrepreneurial risk and global dynamics." *Journal of Economic Theory* 146 (3):863–896.

Avery, Robert, Kenneth Brevoort, and Glenn Canner. 2011. "The Mortgage Market in 2010: Highlights from the Data Reported under the Home Mortgage Disclosure Act." Tech. rep.

Bacchetta, Philippe and Kenza Benhima. 2015. "THE DEMAND FOR LIQUID ASSETS, CORPORATE SAVING, AND INTERNATIONAL CAPITAL FLOWS." *Journal of the European Economic Association* 13 (6):1101–1135. URL http://dx.doi.org/10.1111/jeea.12132.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis. 2015. "Measuring Economic Policy Uncertainty." CEP Discussion Papers dp1379, Centre for Economic Performance, LSE. URL https://ideas.repec.org/p/cep/cepdps/dp1379.html.

Barlevy, Gadi and Jonas D. M. Fisher. 2012. "Mortgage choices and housing speculation." Unpublished working paper. URL https://ideas.repec.org/p/fip/fedhwp/wp-2010-12.html.

Bayer, Patrick, Christopher Geissler, Kyle Mangum, and James W. Roberts. 2011. "Speculators and Middlemen: The Strategy and Performance of Investors in the Housing Market." NBER Working Papers 16784, National Bureau of Economic Research, Inc. URL `https://ideas.repec.org/p/nbr/nberwo/16784.html`.

Bernanke, Ben S. and Cara S. Lown. 1991. "The Credit Crunch." *Brookings Papers on Economic Activity* 22 (2):205–248. URL `https://ideas.repec.org/a/bin/bpeajo/v22y1991i1991-2p205-248.html`.

Berrospide, Jose M., Lamont K. Black, and William R. Keeton. 2016. "The Cross-Market Spillover of Economic Shocks through Multimarket Banks." *Journal of Money, Credit and Banking* 48 (5):957–988. URL `http://dx.doi.org/10.1111/jmcb.12323`.

Bhutta, Neil. 2015. "The ins and outs of mortgage debt during the housing boom and bust." *Journal of Monetary Economics* 76 (C):284–298. URL `https://ideas.repec.org/a/eee/moneco/v76y2015icp284-298.html`.

Bhutta, Neil and Daniel R. Ringo. 2014. "The 2013 Home Mortgage Disclosure Act Data." *Federal Reserve Bulletin* (Nov):A107–A146.

Bloom, Nicholas. 2014. "Fluctuations in Uncertainty." *Journal of Economic Perspectives* 28 (2):153–76. URL `https://ideas.repec.org/a/aea/jecper/v28y2014i2p153-76.html`.

Boldrin, Michele, Carlos Garriga, Adrian Peralta-Alva, and Juan M. Sanchez. 2016. "Reconstructing the great recession." Working Papers 2013-006, Federal

Reserve Bank of St. Louis. URL `https://ideas.repec.org/p/fip/fedlwp/2013-006.html`.

Brunnermeier, Markus K. 2008. "Deciphering the Liquidity and Credit Crunch 2007-08." NBER Working Papers 14612, National Bureau of Economic Research, Inc. URL `https://ideas.repec.org/p/nbr/nberwo/14612.html`.

Buera, Francisco J. and Yongseok Shin. 2017. "Productivity Growth and Capital Flows: The Dynamics of Reforms." *American Economic Journal: Macroeconomics* 9 (3):147–185. URL `https://ideas.repec.org/a/aea/aejmac/v9y2017i3p147-85.html`.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 2016. "Understanding Booms and Busts in Housing Markets." *Journal of Political Economy* 124 (4):1088–1147. URL `https://ideas.repec.org/a/ucp/jpolec/doi10.1086-686732.html`.

Caballero, Ricardo J., Emmanuel Farhi, and Pierre-Olivier Gourinchas. 2008. "An Equilibrium Model of "Global Imbalances" and Low Interest Rates." *American Economic Review* 98 (1):358–393. URL `https://ideas.repec.org/a/aea/aecrev/v98y2008i1p358-93.html`.

Calem, Paul, Francisco Covas, and Jason Wu. 2013. "The Impact of the 2007 Liquidity Shock on Bank Jumbo Mortgage Lending." *Journal of Money, Credit and Banking* 45:59–91. URL `https://ideas.repec.org/a/mcb/jmoncb/v45y2013ip59-91.html`.

Carroll, Christopher D and Olivier Jeanne. 2015. “A tractable model of precautionary reserves, net foreign assets, or sovereign wealth funds.” Tech. rep., National Bureau of Economic Research.

Carroll, Christopher D and Miles S Kimball. 1996. “On the Concavity of the Consumption Function.” *Econometrica* 64 (4):981–92. URL `https://ideas.repec.org/a/ecm/emetrp/v64y1996i4p981-92.html`.

Carroll, Christopher D, Jody Overland, and David N Weil. 2000. “Saving and growth with habit formation.” *American Economic Review* :341–355.

Cetorelli, Nicola and Linda S Goldberg. 2011. “Global Banks and International Shock Transmission: Evidence from the Crisis.” *IMF Economic Review* 59 (1):41–76. URL `http://ideas.repec.org/a/pal/imfecr/v59y2011i1p41-76.html`.

Chamon, Marcos, Kai Liu, and Eswar Prasad. 2013. “Income uncertainty and household savings in China.” *Journal of Development Economics* 105 (C):164–177. URL `https://ideas.repec.org/a/eee/deveco/v105y2013icp164-177.html`.

Chan, Kam Wing. 2013. “China: internal migration.” *The Encyclopedia of Global Human Migration* .

Chen, Brian S., Samuel G. Hanson, and Jeremy C. Stein. 2017. “The Decline of Big-Bank Lending to Small Business: Dynamic Impacts on Local Credit and Labor Markets.” NBER Working Papers 23843, National Bureau of Economic Research, Inc. URL `https://ideas.repec.org/p/nbr/nberwo/23843.html`.

Chinco, Alex and Christopher Mayer. 2016. "Misinformed Speculators and Mispricing in the Housing Market." *The Review of Financial Studies* 29 (2):486–522. URL +http://dx.doi.org/10.1093/rfs/hhv061.

Chodorow-Reich, Gabriel. 2014. "The Employment Effects of Credit Market Disruptions: Firm-level Evidence from the 2008-9 Financial Crisis." *The Quarterly Journal of Economics* 129 (1):1–59. URL http://ideas.repec.org/a/oup/qjecon/v129y2014i1p1-59.html.

Coeurdacier, Nicolas, Stphane Guibaud, and Keyu Jin. 2015. "Credit Constraints and Growth in a Global Economy." *American Economic Review* 105 (9):2838–81. URL http://www.aeaweb.org/articles?id=10.1257/aer.20130549.

Collinson, Mark, Stephen Tollman, Kathleen Kahn, and Samuel Clark. 2006. "Highly prevalent circular migration: Households, mobility and economic status in rural South Africa." .

Cornett, Marcia Millon, Jamie John McNutt, Philip E. Strahan, and Hassan Tehranian. 2011. "Liquidity risk management and credit supply in the financial crisis." *Journal of Financial Economics* 101 (2):297–312. URL http://ideas.repec.org/a/eee/jfinec/v101y2011i2p297-312.html.

Dagher, Jihad and Kazim Kazimov. 2012. "Banks' Liability Structure and Mortgage Lending During the Financial Crisis." IMF Working Papers 12/155, International Monetary Fund. URL https://ideas.repec.org/p/imf/imfwpa/12-155.html.

DeFusco, Anthony A., Charles G. Nathanson, and Eric Zwick. 2017. "Selective Dynamics of Prices and Volume." NBER Working Papers 23449, National Bureau of Economic Research, Inc.

Dell′Ariccia, Giovanni, Deniz Igan, and Luc Laeven. 2012. "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market." *Journal of Money, Credit and Banking* 44:367–384. URL `https://ideas.repec.org/a/mcb/jmoncb/v44y2012ip367-384.html`.

Demyanyk, Yuliya and Otto Van Hemert. 2011. "Understanding the Subprime Mortgage Crisis." *Review of Financial Studies* 24 (6):1848–1880. URL `https://ideas.repec.org/a/oup/rfinst/v24y2011i6p1848-1880.html`.

Deshingkar, Priya and Shaheen Akter. 2009. "Migration and Human Development in India." Human Development Research Papers (2009 to present) HDRP-2009-13, Human Development Report Office (HDRO), United Nations Development Programme (UNDP). URL `https://ideas.repec.org/p/hdr/papers/hdrp-2009-13.html`.

DiMaggio, Marco and Amir Kermani. 2014. "Credit-Induced Boom and Bust." Tech. rep.

Duygan-Bump, Burcu, Alexey Levkov, and Judit Montoriol-Garriga. 2015. "Financing constraints and unemployment: Evidence from the Great Recession." *Journal of Monetary Economics* 75 (C):89–105. URL `https://ideas.repec.org/a/eee/moneco/v75y2015icp89-105.html`.

Eggertsson, Gauti B. and Paul Krugman. 2012. "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach." *The Quarterly Journal* *of Economics* 127 (3):1469–1513. URL http://ideas.repec.org/a/oup/qjecon/v127y2012i3p1469-1513.html.

Elsby, Michael W. L., Bart Hobijn, and Aysegul Sahin. 2010. "The Labor Market in the Great Recession." *Brookings Papers on Economic Activity* 41 (1 (Spring):1–69. URL https://ideas.repec.org/a/bin/bpeajo/v41y2010i2010-01p1-69.html.

Elul, Ronel and Sebastian Tilson. 2015. "Owner occupancy fraud and mortgage performance." Working Papers 15-45, Federal Reserve Bank of Philadelphia. URL https://ideas.repec.org/p/fip/fedpwp/15-45.html.

Favara, Giovanni and Jean Imbs. 2015. "Credit Supply and the Price of Housing." *American Economic Review* 105 (3):958–92. URL https://ideas.repec.org/a/aea/aecrev/v105y2015i3p958-92.html.

Foote, Christopher L., Kristopher S. Gerardi, Lorenz Goette, and Paul S. Willen. 2008. "Subprime facts: what (we think) we know about the subprime crisis and what we don't." Unpublished working paper.

Foote, Christopher L., Lara Loewenstein, and Paul S. Willen. 2016. "Cross-Sectional Patterns of Mortgage Debt during the Housing Boom: Evidence and Implications." NBER Working Papers 22985, National Bureau of Economic Research, Inc. URL https://ideas.repec.org/p/nbr/nberwo/22985.html.

Frame, W. Scott, Andreas Fuster, Joseph Tracy, and James Vickery. 2015. "The rescue of Fannie Mae and Freddie Mac." Staff Reports 719, Federal Reserve

Bank of New York. URL https://ideas.repec.org/p/fip/fednsr/719.html.

Fuster, Andreas, Laurie Goodman, David O. Lucca, Laurel Madar, Linsey Molloy, and Paul S. Willen. 2013. "The rising gap between primary and secondary mortgage rates." *Economic Policy Review* (Dec):17–39. URL https://ideas.repec.org/a/fip/fednep/00002.html.

Gao, Zhenyu, Michael Sockin, and Wei Xiong. 2017. "Economic Consequences of Housing Speculation." Unpublished working paper.

García, Daniel. 2017. "Declines in Mortgage Supply and Employment in the Great Recession." Unpublished working paper.

Garriga, Carlos, Rodolfo E. Manuelli, and Adrian Peralta-Alva. 2012. "A model of price swings in the housing market." Working Papers 2012-022, Federal Reserve Bank of St. Louis. URL https://ideas.repec.org/p/fip/fedlwp/2012-022.html.

Gerardi, Kristopher S., Adam Hale Shapiro, and Paul S. Willen. 2008. "Subprime outcomes: risky mortgages, homeownership experiences, and foreclosures." Working Papers 07-15, Federal Reserve Bank of Boston. URL https://ideas.repec.org/p/fip/fedbwp/07-15.html.

Gete, Pedro and Michael Reher. 2016. "Systemic Banks, Mortgage Supply and Housing Rents." Tech. rep.

Giroud, Xavier and Holger M. Mueller. 2015. "Firm Leverage and Unemployment during the Great Recession." NBER Working Papers 21076, National

Bureau of Economic Research, Inc. URL `http://ideas.repec.org/p/nbr/nberwo/21076.html`.

Glancy, David. 2015. "Housing Bust, Bank Lending & Employment: Evidence from Multimarket Banks." Working paper, JMP Brown University.

Gorton, Gary and Andrew Metrick. 2012. "Securitized banking and the run on repo." *Journal of Financial Economics* 104 (3):425–451. URL `https://ideas.repec.org/a/eee/jfinec/v104y2012i3p425-451.html`.

Gourinchas, Pierre-Olivier and Olivier Jeanne. 2013. "Capital Flows to Developing Countries: The Allocation Puzzle." *Review of Economic Studies* 80 (4):1484–1515. URL `http://ideas.repec.org/a/oup/restud/v80y2013i4p1484-1515.html`.

Gourinchas, Pierre-Olivier and Hélène Rey. 2013. "External Adjustment, Global Imbalances and Valuation Effects." Tech. rep., National Bureau of Economic Research.

Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nguyen. 2015. "Do Credit Market Shocks affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and Normal Economic Times." NBER Working Papers 20704, National Bureau of Economic Research, Inc. URL `http://ideas.repec.org/p/nbr/nberwo/20704.html`.

Gropp, Reint, John Krainer, and Elizabeth Laderman. 2014. "Did consumers want less debt? consumer credit demand versus supply in the wake of the 2008-2009 financial crisis." Working Paper Series 2014-8, Federal Reserve

Bank of San Francisco. URL https://ideas.repec.org/p/fip/fedfwp/2014-08.html.

Guerrieri, Veronica and Guido Lorenzoni. 2011. “Credit Crises, Precautionary Savings, and the Liquidity Trap.” NBER Working Papers 17583, National Bureau of Economic Research, Inc. URL http://ideas.repec.org/p/nbr/nberwo/17583.html.

Haas, Ralph and Iman Lelyveld. 2014. “Multinational Banks and the Global Financial Crisis: Weathering the Perfect Storm?” *Journal of Money, Credit and Banking* 46 (s1):333–364. URL http://ideas.repec.org/a/wly/jmoncb/v46y2014is1p333-364.html.

Haltenhof, Samuel, Seung Jung Lee, and Viktors Stebunovs. 2014. “The credit crunch and fall in employment during the Great Recession.” *Journal of Economic Dynamics and Control* 43 (C):31–57. URL https://ideas.repec.org/a/eee/dyncon/v43y2014icp31-57.html.

Hare, Denise. 1999. “Pushversus pullfactors in migration outflows and returns: Determinants of migration status and spell duration among China’s rural population.” *The Journal of Development Studies* 35 (3):45–72.

Haughwout, Andrew F., Donghoon Lee, Joseph Tracy, and Wilbert van der Klaauw. 2011. “Real estate investors, the leverage cycle, and the housing market crisis.” Staff Reports 514, Federal Reserve Bank of New York. URL https://ideas.repec.org/p/fip/fednsr/514.html.

Henderson, Vernon. 2003. “The Urbanization Process and Economic Growth:

The So-What Question." *Journal of Economic Growth* 8 (1):47–71. URL https://ideas.repec.org/a/kap/jecgro/v8y2003i1p47-71.html.

Irani, Rustom M. and Ralf R. Meisenzahl. 2015. "Loan Sales and Bank Liquidity Risk Management: Evidence from a U.S. Credit Register." Tech. rep.

Ivashina, Victoria and David Scharfstein. 2010. "Bank lending during the financial crisis of 2008." *Journal of Financial Economics* 97 (3):319–338. URL http://ideas.repec.org/a/eee/jfinec/v97y2010i3p319-338.html.

Jaeger, David A., Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, and Holger Bonin. 2010. "Direct Evidence on Risk Attitudes and Migration." *The Review of Economics and Statistics* 92 (3):684–689. URL https://ideas.repec.org/a/tpr/restat/v92y2010i3p684-689.html.

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti. 2013. "The Effects of the Saving and Banking Glut on the U.S. Economy." NBER Working Papers 19635, National Bureau of Economic Research, Inc. URL https://ideas.repec.org/p/nbr/nberwo/19635.html.

Justiniano, Alejandro, Giorgio E Primiceri, and Andrea Tambalotti. 2015. "Credit Supply and the Housing Boom." CEPR Discussion Papers 10358, C.E.P.R. Discussion Papers. URL https://ideas.repec.org/p/cpr/ceprdp/10358.html.

Kacperczyk, Marcin and Schnabl. 2010. "When Safe Proved Risky: Commercial Paper during the Financial Crisis of 2007-2009." *Journal of Economic Perspectives* 24 (1):29–50. URL https://ideas.repec.org/a/aea/jecper/v24y2010i1p29-50.html.

Kaplan, Greg, Kurt Mitman, and Giovanni L. Violante. 2017. "The Housing Boom and Bust: Model Meets Evidence." NBER Working Papers 23694, National Bureau of Economic Research, Inc. URL `https://ideas.repec.org/p/nbr/nberwo/23694.html`.

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. 2010. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans." *The Quarterly Journal of Economics* 125 (1):307–362. URL `https://ideas.repec.org/a/oup/qjecon/v125y2010i1p307-362..html`.

Khwaja, Asim Ijaz and Atif Mian. 2008. "Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market." *American Economic Review* 98 (4):1413–42. URL `http://ideas.repec.org/a/aea/aecrev/v98y2008i4p1413-42.html`.

Kung, Edward. 2015. "The Effect of Credit Availability on House Prices: Evidence from the Economic Stimulus Act of 2008." Tech. rep., Working Paper.

Lacko, James M. and Janis K. Pappalardo. 2010. "The Failure and Promise of Mandated Consumer Mortgage Disclosures: Evidence from Qualitative Interviews and a Controlled Experiment with Mortgage Borrowers." *American Economic Review* 100 (2):516–521. URL `https://ideas.repec.org/a/aea/aecrev/v100y2010i2p516-21.html`.

Lucas, Robert E.B. 1993. "Internal migration in developing countries." In *Handbook of Population and Family Economics*, *Handbook of Population and Family Economics*, vol. 1, edited by M. R. Rosenzweig and O. Stark, chap. 13.

Elsevier, 721–798. URL https://ideas.repec.org/h/eee/popchp/1-13.html.

Manski, Charles F. 1993. “Identification of Endogenous Social Effects: The Reflection Problem.” *Review of Economic Studies* 60 (3):531–542. URL https://ideas.repec.org/a/oup/restud/v60y1993i3p531-542..html.

Mian, Atif, Kamalesh Rao, and Amir Sufi. 2013. “Household Balance Sheets, Consumption, and the Economic Slump.” *The Quarterly Journal of Economics* 128 (4):1687–1726. URL http://ideas.repec.org/a/oup/qjecon/v128y2013i4p1687-1726.html.

Mian, Atif and Amir Sufi. 2009. “The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis*.” *The Quarterly Journal of Economics* 124 (4):1449. URL +http://dx.doi.org/10.1162/qjec.2009.124.4.1449.

———. 2011. “House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis.” *American Economic Review* 101 (5):2132–56. URL http://ideas.repec.org/a/aea/aecrev/v101y2011i5p2132-56.html.

———. 2014. “What Explains the 20072009 Drop in Employment?” *Econometrica* 82:2197–2223. URL https://ideas.repec.org/a/wly/emetrp/v82y2014ip2197-2223.html.

Mian, Atif R. and Amir Sufi. 2015. “Fraudulent Income Overstatement on Mortgage Applications during the Credit Expansion of 2002 to 2005.” NBER Working Papers 20947, National Bureau of Economic Research, Inc. URL https://ideas.repec.org/p/nbr/nberwo/20947.html.

Midrigan, Virgiliu and Thomas Philippon. 2016. "Household Leverage and the Recession." CEPR Discussion Papers 11407, C.E.P.R. Discussion Papers. URL https://ideas.repec.org/p/cpr/ceprdp/11407.html.

Mills, James, Raven Molloy, and Rebecca Zarutskie. 2017. "Large-Scale Buy-to-Rent Investors in the Single-Family Housing Market: The Emergence of a New Asset Class." *Real Estate Economics* :n/a–n/aURL http://dx.doi.org/10.1111/1540-6229.12189.

Modigliani, Franco. 1970. "The life cycle hypothesis of saving and intercountry differences in the saving ratio." *Induction, growth and trade* :197–225.

Mondragon, John. 2014. "Household Credit and Employment in the Great Recession." Unpublished working paper, UC Berkeley.

———. 2018. "Household Credit and Employment in the Great Recession." Unpublished working paper, Kellogg School of Management. URL https://sites.google.com/site/johnnelsonmondragon/.

Nadauld, Taylor D. and Shane M. Sherlund. 2009. "The Role of the Securitization Process in the Expansion of Subprime Credit." Working Paper Series 2009-9, Ohio State University, Charles A. Dice Center for Research in Financial Economics. URL https://ideas.repec.org/p/ecl/ohidic/2009-9.html.

Newland, Kathleen. 2009. "Circular Migration and Human Development." Human Development Research Papers (2009 to present) HDRP-2009-42, Human Development Report Office (HDRO), United Nations Development Programme (UNDP). URL https://ideas.repec.org/p/hdr/papers/hdrp-

2009-42.html.

Nguyen, Hoai-Luu. 2014. "Do Bank Branches Still Matter? The Effect of Closings on Local Economic Outcomes." Working paper, MIT.

Nieuwerburgh, Stijn Van and Jack Favilukis. 2017. "Out-of-town Home Buyers and City Welfare." 2017 Meeting Papers 486, Society for Economic Dynamics. URL https://ideas.repec.org/p/red/sed017/486.html.

Passmore, Wayne and Shane M. Sherlund. 2016. "Government-Backed Mortgage Insurance, Financial Crisis, and the Recovery from the Great Recession." Tech. rep.

Passmore, Wayne, Shane M. Sherlund, and Gillian Burgess. 2005. "The Effect of Housing Government-Sponsored Enterprises on Mortgage Rates." *Real Estate Economics* 33 (3):427–463. URL https://ideas.repec.org/a/bla/reesec/v33y2005i3p427-463.html.

Piazzesi, Monika and Martin Schneider. 2009. "Momentum Traders in the Housing Market: Survey Evidence and a Search Model." *American Economic Review* 99 (2):406–11. URL http://www.aeaweb.org/articles?id=10.1257/aer.99.2.406.

Piketty, Thomas and Gabriel Zucman. 2013. "Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010." CEPR Discussion Papers 9588, C.E.P.R. Discussion Papers. URL http://ideas.repec.org/p/cpr/ceprdp/9588.html.

Piskorski, Tomasz, Amit Seru, and James Witkin. 2015. "Asset Quality Misrepresentation by Financial Intermediaries: Evidence from the RMBS Market." *The Journal of Finance* 70 (6):2635–2678. URL `http://dx.doi.org/10.1111/jofi.12271`.

Quadrini, Vincenzo, E Mendoza, and V Rios-Rull. 2009. "Financial integration, financial deepness and global imbalances." *Journal of Political Economy* 117 (3).

Ramcharan, Rodney, Skander Van den Heuvel, and Stephane Verani. 2013. "From Wall Street to main street: the impact of the financial crisis on consumer credit supply." Tech. rep.

Rognlie, Matthew, Andrei Shleifer, and Alp Simsek. Forthcoming. "Investment Hangover and the Great Recession." *American Economic Journal: Macroeconomics* .

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *The Quarterly Journal of Economics* 125 (3):1253–1296. URL `https://ideas.repec.org/a/oup/qjecon/v125y2010i3p1253-1296..html`.

Sandri, Damiano. 2014. "Growth and Capital Flows with Risky Entrepreneurship." *American Economic Journal: Macroeconomics* 6 (3):102–123. URL `https://ideas.repec.org/a/aea/aejmac/v6y2014i3p102-23.html`.

Schnabl, Phillip. 2012. "The International Transmission of Bank Liquidity Shocks: Evidence from an Emerging Market." *The Journal of Finance* 67 (3):897–932. URL `http://dx.doi.org/10.1111/j.1540-6261.2012.01737.x`.

Shin, Hyun Song. 2012. "Global Banking Glut and Loan Risk Premium." *IMF Economic Review* 60 (2):155–192. URL http://ideas.repec.org/a/pal/imfecr/v60y2012i2p155-192.html.

Skeldon, Ronald. 2012. "Going Round in Circles: Circular Migration, Poverty Alleviation and Marginality." *International Migration* 50 (3):43–60. URL http://dx.doi.org/10.1111/j.1468-2435.2012.00751.x.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing like china." *The American Economic Review* 101 (1):196–233.

Stark, Oded. 1993. "The migration of labor : Oded Stark, (Basil Blackwell, Oxford, 1991) pp. x+406." *Regional Science and Urban Economics* 23 (3):453–457. URL https://ideas.repec.org/a/eee/regeco/v23y1993i3p453-457.html.

Stock, James H. and Motohiro Yogo. 2002. "Testing for Weak Instruments in Linear IV Regression." NBER Technical Working Papers 0284, National Bureau of Economic Research, Inc. URL https://ideas.repec.org/p/nbr/nberte/0284.html.

Taylor, Guy. 2011. "China's Floating Migrants." Tech. rep.

Thom, Kevin. 2010. "Repeated Circular Migration: Theory and Evidence from Undocumented Migrants." Unpublished working paper.

Vojtech, Cindy, Benjamin Kay, and John Driscoll. 2016. "The Real Consequences of Bank Mortgage Lending Standards." Tech. rep.

Williamson, Jeffrey G. 1988. "Migration and urbanization." In *Handbook of Development Economics*, *Handbook of Development Economics*, vol. 1, edited by Hollis Chenery and T.N. Srinivasan, chap. 11. Elsevier, 425–465. URL `https://ideas.repec.org/h/eee/devchp/1-11.html`.

Wing Chan, Kam and Will Buckingham. 2008. "Is China abolishing the hukou system?" *The China Quarterly* 195:582–606.

Woodward, Susan E. and Robert E. Hall. 2012. "Diagnosing Consumer Confusion and Sub-optimal Shopping Effort: Theory and Mortgage-Market Evidence." *American Economic Review* 102 (7):3249–76. URL `http://www.aeaweb.org/articles?id=10.1257/aer.102.7.3249`.

Zhan, Shaohua. 2005. "Rural labour migration in China: Challenges for policies." *MOST-2 Policy. France: UNESCO Papers* 10.

Zhao, Zhong. 2005. "Migration, labor market flexibility, and wage determination in China: A review." *The Developing Economies* 43 (2):285–312.

Curriculum Vitae

Daniel García grew up in Tegucigalpa, Honduras. He attended Grinnell College in the fall of 2008, where he earned his B.A. in economics in 2012. Later that year, he joined the Johns Hopkins University economics Ph.D. program. During graduate school, he received the Carl Christ Award and was a dissertation fellow at the Board of Governors of the Federal Reserve System, and the Federal Reserve Bank of Chicago.

In September 2018 he will join the Board of Governors of the Federal Reserve System in Washington, D.C.